12-1-2014

Teacher Perceptions of the Impact of an Evaluation System on Classroom Instructional Practices

Erika Donahue

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

TEACHER PERCEPTIONS OF THE IMPACT OF AN EVALUATION SYSTEM ON CLASSROOM INSTRUCTIONAL PRACTICES

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Education

Erika Donahue

College of Education and Behavioral Sciences
Department of Leadership, Policy and Development
Program of Educational Leadership and Policy Studies

December 2014

This Dissertation by: Erika Donahue

Entitled: *Teacher Perceptions of the Impact of an Evaluation System on Classroom Instructional Practices*

has been approved as meeting the requirement for the Degree of Doctor of Education in College of Education and Behavioral Sciences in Department of Leadership, Policy and Development, Program of Educational Leadership and Policy Studies

Accepted by the Doctoral Committee

_____
Anthony Armenta, Ed.D., Co-Research Advisor

_____
Linda R. Vogel, Ph.D., Co-Research Advisor

_____
Spencer C. Weiler, Ph.D., Committee Member

_____
Jody K. Lawrence, Ph.D., Faculty Representative

Date of Dissertation Defense _____

Accepted by the Graduate School

_____
Linda L. Black, Ed.D.
Dean of the Graduate School and International Admissions

ABSTRACT

Donahue, Erika. *Teacher Perceptions of the Impact of an Evaluation System on Classroom Instructional Practices.* Published Doctor of Education dissertation, University of Northern Colorado, 2014.

Teacher evaluations systems are becoming a method for measuring teacher effectiveness as well as a method for improving teacher instructional practice. Cycles of observation, feedback, and professional development are essential for improving the practice of teaching and the education of students overall. The researcher synthesized the experiences of 30 teachers regarding their perceptions of the evaluation system in which they conduct their professional practice.

The study sample included teacher representation from most grades and levels. Kindergarten through grade 12 teachers, special educators, physical education teachers, art teachers, language specialists, elementary teachers, secondary teachers, and teachers of literacy, math, science, and social studies were represented as well as teachers in coaching roles.

Five major themes emerged from this study and identified the ways teachers perceived an existing system of evaluation to influence their classroom instructional practices. The effect of feedback, quality of relationships, effect of the professional practices rubric, effect of modeling, and the effect of personal integrity and self-reflection were identified as mechanisms that enabled teachers to benefit from the existing system.

These themes revealed the complexity of the system in which multiple mechanisms must work in coordination in order for teachers to realize the benefits.

ACKNOWLEDGEMENTS

Thank you to my research advisors, Anthony Armenta and Linda Vogel, for the ongoing support and direction provided throughout the duration of this study. Each of you contributed something unique to my study through your individual areas of interest. Dr. Armenta, thank you for your attention to the detail in my writing and for your belief in me. Dr. Vogel, thank you for sharing my interest in teacher evaluation systems, and agreeing that teacher perceptions are worthy of study. To my other committee members, Spencer Weiler and Jody Lawrence, thank you for your contributions to the design of this study and for asking the tough questions--the ones that made me wonder and reconsider the perspectives to which I held tightly. It is my hope we will all have an opportunity to work together in the future.

An extended thank you is due to the individuals who lent an expert hand in taking this dissertation to completion. Thank you to transcription expert, Becky Hendren, and dissertation editor extraordinaire, Connie Beard. Thank you for your typing wizardry!

Additionally, I would like to thank the teachers who participated in this study. You gave up time on your weekends, school vacations, and after school planning time to meet with me and I appreciate that. I know how difficult it is to find even a few minutes for yourself when you are a teacher and yet you gave those few minutes to me. Thank you!

I would also like to thank my family, Patty, Barry, Alison, Marty, and Kristle for their support and interest in my work. Your belief that I would survive this challenge gave me the energy to persevere.

Finally, I would like to acknowledge my husband, Steve, and my daughter, Quinn. The path to a doctoral degree is not easy on a family. I appreciate the ability Steve and Quinn demonstrated to get along on their own as I made the weekly six hour trek to and from Greeley. I appreciate their waning interest in home cooking as that is something that disappeared as the rigors of doctoral work took hold. And I appreciate Steve and Quinn for allowing me to bring a computer, or draft of whatever current paper I was working on, to every family vacation over the past four years. Your belief that I would conquer this challenge helped me stay the course even when other things seemed so much more important. No one could be luckier than I to have a husband and daughter like you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

More than two decades of government programs have attempted to address challenges within the American public school system (Darling-Hammond, 2010). Negative perceptions surrounding the quality of education received by students in public schools have led to interventions ranging from new standards and tests to redesigned schools, new curricula, and new governance models (p. 2). All have unfortunately missed the mark in school improvement. *What Matters Most: Teaching for America's Future*, a 1996 publication by the National Commission on Teaching and America's Future, identified teachers as the most powerful school-based intervention. Weisberg, Sexton, Mulhern, and Keeling (2009), Papay (2012), and Kane and Cantrell (2012) confirmed this finding and highlighted the importance of teacher effectiveness on student achievement. An implication of these findings is an overall shift in the focus of school reform to improve the practice of teaching. Every aspect of school reform depends on highly skilled teachers for its success (Darling-Hammond, 2010).

As the demands for teacher quality have escalated into a national preoccupation, the public has exercised its right to demand forceful and effective leadership on issues of school quality and teacher effectiveness (Stand Organization, 2010). The emphasis on school improvement has become narrowly focused on measuring teacher effectiveness and improving teaching practices. Fueled in part by the U.S. Department of Education's

(2009) *Race to the Top* program, a massive effort to overhaul teacher evaluation is underway in states and school districts. The aim is to ensure that evaluations provide a better indication of "teaching effectiveness" or the extent to which teachers can and do contribute to student learning and then act on that information to enhance teaching and learning (Jerald, 2012, p. 1). Teacher evaluation systems are developing a reputation as a "catalyst for teacher and school improvement" (Toch & Rothman, 2008, p. 32).

Expanding government attention has initiated legislative and policy decisions with a laser-like focus on teacher quality and teacher effectiveness. Colorado has made an educationally defining decision in the passage of Senate Bill 191--Ensuring Quality Instruction Through Educator Effectiveness, also known as the Educator Effectiveness Act (2010). Districts throughout the state of Colorado are faced with the challenges of measuring and reporting teacher effectiveness while simultaneously trying to improve teacher instructional practices as implementation of the new state model of teacher evaluation begins (Colorado Legacy Foundation, 2013).

Problem Statement

The troubled state of teacher evaluation is a glaring and largely neglected problem in public education, one with consequences that extend far beyond the performance pay debate (Toch & Rothman, 2008) and one that has persisted since 1985 when Albert Shanker, the president of the American Federation of Teachers, made a compelling case for union support of rigorous evaluations:

> We don't have the right to be called professionals, and we will never convince the public that we are- unless we are prepared honestly to decide what constitutes competence in our profession and apply those definitions to ourselves and our colleagues. (as cited in Toch & Rothman, 2008)

The most effective way to strengthen both teaching and learning is to put in place evaluation systems that are not just a stamp of approval or disapproval but also a means of improvement (Phillips & Weingarten, 2013).

The importance of implementing a comprehensive evaluation system that meets the criteria of the Colorado Model for Teacher Effectiveness (Colorado Department of Education [CDE], 2014) goes far beyond policy mandates identified by the law. If teacher evaluation systems truly are the strategy for large-scale school improvement, systems must provide a definition of effective teaching as well as the tools to get there. For many school districts in Colorado, the state model rubric for teacher evaluation is the method that will be used to measure, label, and support teacher growth. However, some school districts have developed their own comprehensive systems for evaluation. Whether using the state model or one developed internally by a school district, the question addressing whether an existing system will provide teachers with what is needed to improve practice while meeting the criteria of teacher effectiveness legislation still remains.

The criteria within the Educator Effectiveness Act (2010) and Colorado teacher quality standards may influence the redesign of evaluation systems currently existing in school districts across the state. One Colorado school district is striving to adjust an existing evaluation system to reflect the teacher quality standards and elements of the state model. It is imperative to study the perceived purpose and effect of an existing evaluation system from the point of view of teachers. Data gleaned from this study might inform decision-making for future evaluation system design.

Research Questions

While many school districts in Colorado work through the details of implementing the new state system for teacher evaluation, this researcher was compelled to examine the effect of an existing system, specific to one Colorado school district, on the improvement of teaching practices. The following research questions guided this investigation:

Q1 How do novice teachers feel the district evaluation system impacts their classroom practice?

Q2 How do experienced teachers feel the district evaluation system impacts their classroom practice?

Purpose of the Study

School reform is about identifying instructional practices that work and replicating them. Teacher quality has been identified as the most powerful school-based intervention, placing teacher evaluation at the forefront of the minds of stakeholders (Kane & Cantrell, 2012; Papay, 2012; Schools Matter, 2010; Weisberg et al., 2009). If school districts seek to produce widespread and systematic improvement in student learning, then efforts to reform teacher evaluation should focus on continued teacher improvement (Heneman & Milanowski, 2003; Kimball, 2002; Odden, 2004; Papay, 2012; Varlas, 2012). A purpose of this study was to identify perceptions held by teachers who were currently observed and evaluated on a particular set of instructional criteria existing within one Colorado school district. Further, the purpose of the study was to understand perceptions of aspects within the existing evaluation system that were viewed as valuable in terms of helping teachers improve their instructional practices.

Studying perceptions of the impact of evaluation from the perspectives of teachers is important since research clearly identified dual purposes for evaluation. Researchers want to examine how evaluation might be used to measure teacher effectiveness but also want to determine methods for using evaluation to improve teacher practices (Papay, 2012). School district personnel in Colorado school districts want to know how to meet the criteria of the Educator Effectiveness Act and its increased accountability for evaluating teachers while providing supports for improving teacher instructional practices (Colorado Legacy Foundation, 2013). The state's mandate for reporting teacher effectiveness at the individual teacher level has increased the urgency for districts to examine existing systems (Colorado Department of Education, 2014). How teachers perceive an evaluation system intended to improve instructional practices might be important for districts to discover. The perception of evaluation practices viewed by teachers and evaluators to improve instruction and the way in which a system of evaluation informs change in teaching practices might lead to increased understanding of how an existing system might be merged with the new state model.

Conceptual Framework

The conceptual framework for this study reflected the idea that teacher evaluation is a means to improving the educator workforce by providing specific, observation-based feedback; follow-through specifically designed to meet teachers' needs at their current level of practice; and opportunities to provide teachers with information to move their practice forward. Emerging evidence from research regarding comprehensive evaluation models showed that teachers' ratings under comprehensive classroom evaluations aligned with their students' test scores (Darling-Hammond, 2000; Toch & Rothman, 2008;

Weglinsky, 2000). Studies are beginning to suggest that teacher instructional practice, as measured by some evaluation systems, can be improved through evaluation as well as support student growth (Kane, Wooten, Taylor & Tyler, 2011; Milanowski, 2004; Taylor & Tyler, 2011; The New Teacher Project [TNTP], 2013). Further, evaluations of teacher performance might predict effectiveness as measured by student achievement (Kimball, White, Milanowski, & Borman, 2004; Milanowski, 2004; Milanowski, Kimball, & Odden, 2005).

Phillips and Weingarten (2013) suggested that teachers should receive regular and timely feedback on their performance, corresponding support to improve, and to receive constructive feedback from evaluators prepared to give skilled feedback that can support teachers' growth. Jerald (2012) offered the idea of "improvin' it," strategies reflecting the view that each individual's measured effectiveness is a "mutable trait that can improve over time" (p. 4). He further suggested providing teachers with support, assistance, and useful feedback to enhance knowledge and practice. In a study of teacher responses to feedback, enabling conditions, and perceptions of fairness in evaluation systems, Kimball (2002) developed a conceptual model that provided a view of how classroom observations might translate into improved teacher practices. Kimball's model put forth the idea that whether or not teachers decide to use the feedback provided through observation and evaluation is influenced by a common, comprehensive understanding of teaching, the quality of feedback, perceived fairness in the system, and enabling conditions in the organization. If evaluation is to be a vehicle to achieve the mission of public schooling (Phillips & Weingarten, 2013) and if teacher evaluations are a potentially powerful lever of school improvement (Toch & Rothman, 2008),

improvements to teacher evaluation systems should focus on setting clear, high expectations for teachers and providing them with useful feedback and support in order to help their students master needed skills (TNTP, 2013).

Definitions

Understanding the complexities of teacher evaluation systems is a challenging task and the varied definitions surrounding common terminology make it difficult for educators to make sense of what they are being asked to do and how they are being measured. For the purpose of this research study, the following definitions serve to provide common language for frequently used terms.

Coach. Synonymous with mentor teacher. The coach observes assigned teachers and provides feedback aligned with teaching standards to improve classroom instruction and student achievement.

Effective teacher. Effective teachers in the state of Colorado have the knowledge, skills, and commitments that ensure equitable learning opportunities and growth for all students. They strive to close achievement gaps and prepare diverse student populations for postsecondary success. Effective teachers facilitate mastery of content and skill development and identify and employ appropriate strategies for students who are not achieving mastery. They also develop in students the skills, interests, and abilities necessary to be lifelong learners as well as for democratic and civic participation. Effective teachers communicate high expectations to students and their families and find ways to engage them in a mutually supportive teaching and learning environment. Because effective teachers understand that the work of ensuring meaningful learning opportunities for all students cannot happen in

isolation, they engage in continuous reflection, on-going learning, and leadership within the profession (Colorado Department of Education, 2011, p. 9)

Evaluation system. System used to define and measure teaching. The system might include a consistent definition of good teaching, a shared understanding of the definition of good teaching, and skilled evaluators (Danielson, 2011).

Evaluation tool. A rubric including a set of criteria defining and describing the important components of the work being evaluated. For the purpose of this research, an evaluation tool defined the criteria for good teaching and measured how well teachers met those criteria.

Evaluator. A principal, assistant principal, or non-administrative employee with annual requisite training and certification to conduct teacher evaluations using the district prescribed evaluation instrument. The person must be certified in the use of evaluation tools and trained in the process of providing written and face-to-face feedback to teachers. In the case under study in this research, there were four professional roles within school buildings that might be held by evaluators: mentor teacher, master teacher, assistant principal, and principal.

Experienced teacher. Certified teacher having taught more than three years.

Master teacher. Certified teacher with responsibility to work with the principal to design and communicate professional development activities for professional learning communities and/or individual teachers that are aligned with content standards and student achievement data. The master teacher coaches teachers on instruction that directly impacts students at the classroom level. An additional description of the master teacher role is provided in Appendix A.

Mentor teacher. Synonymous with coach. The coach observes assigned teachers and provides feedback aligned with teaching standards to improve classroom instruction and student achievement. An additional description of the mentor teacher role is provided in Appendix A.

Novice teacher. Certified teacher having taught three or less years.

Observation. An opportunity for data collection regarding teacher instruction and student responses to instruction. Data are collected, related to a rubric, and later used to inform a post-conference and evaluative scoring.

Observer. A principal, assistant principal, or non-administrative employee with annual requisite training and certification to conduct teacher observations using the district prescribed observation instrument. The person should be certified in the use of evaluation tools and trained in the process of providing written and face-to-face feedback to teachers. In the district under study in the research, four professional roles within school buildings might be held by evaluators: mentor teacher, master teacher, assistant principal, principal.

Pre-conference. An opportunity, prior to a formal lesson observation, for a teacher to explain to an evaluator how a lesson was planned and how the lesson would be experienced by students.

Post-conference. A face-to-face conference between an evaluator and a teacher regarding the observed lesson. During the post conference, an area of strength is identified and a strategy to continue enhancing the area of strength is shared. An area of weakness or comparative weakness is identified and feedback including a suggestion and modeling for improvement are discussed.

Self-reflection. A document that supports the teacher's own self-reflection and provides the evaluator with evidence of areas the teacher might wish to discuss more in-depth.

Significance of the Study

It is not enough to measure teacher performance and label it when evaluation has the potential to improve it (Danielson, 2011). Toch (2008) suggested that teacher evaluations are at the very center of the education enterprise and can be catalysts for teacher and school improvement. Commentary regarding the professional practices related to the Educator Effectiveness Act (2010) specifically addressed using the state evaluation system to provide teachers with support and feedback leading to continuous improvement and professional growth and improving results for all children.

The irony of all the work that has gone into teacher evaluation is that few researchers have asked teachers to describe their perceptions of how teacher evaluation improved their practice. Evidence of teacher evaluations as sources of information on the types of instructional behavior that affect student learning might be growing as reported in Milanowski (2004) but existing research provided little insight into how teachers perceived the evaluation system to improve their practice, placing it in a position for potential research. Research has not consistently identified a connection between evaluation and substantial improvements, nor expanded learning outcomes for students (Donaldson, 2009), nor have evaluation practices been consistently reported to improve practices or accurately represent what happens in the classroom (Peterson, 2000). Very little is known about how information regarding teaching practices or the experience of being evaluated might change teacher effort and effectiveness (Taylor & Tyler, 2012).

The idea is that evaluation will improve teacher effectiveness and therefore boost achievement but research has not yet demonstrated the mechanisms at play to make this happen.

Conclusion

In this chapter, the purpose and significance of this study regarding teacher perceptions of the impact of an evaluation system on classroom practices were briefly outlined. While research increasingly identifies the classroom teacher as the most impactful school-based influence, little research has shown that teacher evaluations support the growth of effective teaching practices. Developers of evaluation systems, evaluators, and practitioners want to understand what aspects of teacher evaluation systems require more focused attention in order promote professional learning and improve the outcome of instructional practices (Danielson, 2011). Chapter II describes the path taken by public education from the national to the local level through an extensive review of the literature focused on the classroom observation aspect of teacher evaluation systems. In Chapter III, the selection of qualitative methodology and case study for this research on how teachers perceive an evaluation system to inform change in their instructional practices are discussed.

CHAPTER II

REVIEW OF THE LITERATURE

Teacher evaluation in the United States has been a hotly contested issue for many years. Existing systems note deficiencies stemming from outmoded evaluative criteria, simplistic evaluative comments, lack of differentiation reflecting veteran teachers' experience and expertise, lack of consistency among evaluators, and one-way, top-down communication (Danielson, 2011). However, research is beginning to identify well-designed and implemented teacher evaluation systems as one effective way to improve student achievement (Donaldson, 2009; Kane et al., 2011; Kimball et al., 2004; Milanowski, 2004; Milanowski et al., 2005; Taylor & Tyler, 2012; TNTP, 2013). Developing systems that accurately measure teachers' instructional abilities and take into account how instruction translates into student achievement has surfaced as a tremendous challenge in this era of teacher reform.

The discussion of teacher evaluation has often focused on systems as sorting mechanisms--a means to identify the lowest performing teachers for selective termination (Taylor & Tyler, 2012). Due to vague district standards, poor evaluation instruments, and overly restrictive bargaining agreements, school districts are likely to employ more under-performing teachers than their evaluation ratings suggest, creating "Lake Wobegon," where all teachers are satisfactory, yet student achievement is not (Donaldson, 2009, p. 9). Weisberg et al. (2009) identified a teacher's effectiveness as the

most important factor for schools in improving student achievement and added that a teacher's effectiveness has not been measured, recorded, or used to inform decision-making in any standardized or meaningful way. *The Widget Effect* describes the tendency of school districts to assume effectiveness is the same from teacher to teacher and is characterized by teacher ratings that are all good or great, lack recognition of excellence, and have inadequate professional development (Weisberg et al., 2009). When an evaluation system overlooks individual strengths and weaknesses, it becomes insulting to teachers and when evaluation systems are apathetic toward instructional effectiveness, they gamble with the lives of students (Weisberg et al., 2009).

Moore-Johnson (n.d.) proclaimed, "The sorry state of teacher evaluation can no longer be ignored because it threatens the future of the teaching profession and, thus, the future of public education" (p. 1). Taylor and Tyler (2012) identified the evaluation of teacher effectiveness as a dominant theme in American reform efforts motivated by the large variation in teacher productivity as measured by the ability to promote student growth. Teacher evaluation in many districts must be reformed to improve instruction and teaching quality (Papay, 2012) because "not only can no one tell you which teachers are effective, they also cannot say which are the least effective or which fall between" (Weisberg et al., 2009. p. 31).

Teacher Effectiveness

Using value-added measurement, Wright, Horn, and Sanders (1997) found that differences in teacher effectiveness were the dominant factor affecting student academic gain and documented the most important school-based factor to affect student learning was the teacher. Students in classrooms of effective teachers after having relatively

ineffective teachers make excellent academic gains but not enough to offset previous evidence of less than expected gains (Wright et al., 1997).

Gordon, Kane, and Staiger (2006) presented similar findings and reported:

> Average students who were assigned a teacher in the bottom quartile lost on average 5 percentile points relative to students with similar baseline scores and demographics. In contrast, the average student assigned to a top quartile teacher gained 5 points relative to students with similar baseline scores and demographics. Therefore, the average difference between being assigned a top quartile or a bottom quartile teacher is ten percentile points. The black and white achievement gap nationally is roughly 34 percentile points. Therefore, if the effects were to accumulate having a top quartile teacher rather than a bottom quartile teacher four years in a row would be enough to close the black and white test score gap. (p. 8)

More than five years later, a 2010 U. S. Department of Education research synthesis found that top performing teachers could make a dramatic difference in the achievement of their students and suggested the impact of being assigned to top-performing teachers year after year was enough to significantly narrow achievement gaps. In all high-performing education systems, significant attention is placed on educator quality as a driver of student success (Glass, 2013). Teacher selectivity, quality, and growth, regardless of culture or context, make a difference (Battelle for Kids, 2012).

Past efforts to reform teacher evaluation have failed because systems did not accurately measure teacher quality (Marzano, 2012). Value-added models themselves are limited by imperfections in the tests, inadequate year-to-year tracking of student progress, and incomplete links between students and teachers, particularly in the case of highly mobile students (Steele, Hamilton, & Stecher, 2010). Researchers argued that evaluation systems failed to discriminate between effective and ineffective teachers and had not aided in developing a highly skilled workforce (Toch & Rothman, 2008; Weisberg et al., 2009). Danielson (2012) pointed out that deficient classroom practices prevent an

individual from being considered a good teacher. Over time, studies of teacher evaluation practices showed little effect on improving teachers or accurately representing what happened in the classroom (Donaldson, 2009; Ellet & Teddlie, 2003; Peterson, 2000).

Colorado Senate Bill 191: Ensuring Quality Education Through Educator Effectiveness

Senate Bill 191 (Educator Effectiveness Act, 2010) was known by a variety of names prior to being passed into law. The Teacher Tenure Bill and the Teacher Effectiveness Bill were common monikers. Both names addressed valid points but the formal title of the law when passed became Ensuring Quality Education Through Educator Effectiveness--in short, the Educator Effectiveness Act (2010). Reforming the practice of tenure, defining teacher effectiveness, and developing parameters for teacher and principal evaluation systems were key components. The spirit of the Educator Effectiveness Act was to use teacher evaluation to provide meaningful feedback to educators so they could continuously improve their practice (Colorado Department of Education, 2011) but knowing that one's tenure is partially tied to the outcome has made teacher evaluation a contentious issue in Colorado's public schools (Meyer, 2010).

A belief inherent in the Educator Effectiveness Act (2010) was a system to evaluate the effectiveness of licensed personnel was crucial to improving the quality of education in Colorado. The Act ensured that one of the purposes of evaluation was to provide a basis for making decisions in the areas of hiring, compensation, promotion, assignment, earning and retaining non-probationary status, and nonrenewal of contracts. It also included requirements that educators would be evaluated based on the impact they had on the growth of their students (Colorado Department of Education, 2011). The idea

of losing one's tenure or being shamed out of the teaching profession based on low evaluation scores were undoubtedly pieces of the legislation that produced stress for the educator workforce.

The Educator Effectiveness Act (2010) initiated a system for teacher evaluation, the Colorado Model for Teacher Effectiveness, founded on the premise that effective teachers would promote growth and achievement for their students. The Colorado Department of Education (2011) *Users Guide* provided the following definition of an effective teacher:

> Effective teachers in the state of Colorado have the knowledge, skills and commitments that ensure equitable learning opportunities and growth for all students. They strive to close achievement gaps and to prepare diverse student populations for postsecondary success. Effective teachers facilitate mastery of content and skill development, and identify and employ appropriate strategies for students who are not achieving mastery. They also develop in students the skills, interests and abilities necessary to be lifelong learners, as well as for democratic and civic participation. Effective teachers communicate high expectations to students and their families and find ways to engage them in a mutually supportive teaching and learning environment. Because effective teachers understand that the work of ensuring meaningful learning opportunities for all students cannot happen in isolation, they engage in continuous reflection, on-going learning and leadership within the profession. (p. 9)

Identifying the practices of effective teachers was critical in the development of this definition and addresses the way instruction might translate into student learning and achievement.

The model for teacher effectiveness embedded within the Educator Effectiveness Act (2010) holds to dual purposes: measuring teachers and developing teachers (Colorado Department of Education, 2013). Each purpose presents different implications since evaluation systems designed for measurement look different from systems designed for development (Marzano, 2012). Odden (2004) and Marzano (2012) offered similar

models of effective evaluation systems designed for teacher growth. Odden proposed a model for a system designed for development guided by (a) clear and specific teaching standards in which multiple forms of data on teacher instructional practices were gathered, (b) trained evaluators scored the data to performance levels according to scoring rubrics, and (c) validation—meaning the higher the evaluation score, the more learning gains will be measured for students in those teachers classrooms. The Marzano model offered three primary characteristics necessary for teacher evaluation systems designed to help teachers improve:

- The system is comprehensive and specific. The model includes all elements that research has identified as associated with student achievement including constructs such as classroom management and teacher-student relationships. Specific means the model identifies classroom strategies and behaviors at a granular level.
- The system includes a developmental scale teachers can use to guide and track their skills development.
- The system acknowledges and rewards growth.

The Colorado Model for Teacher Effectiveness (Educator Effectiveness Act, 2010) attempted to meet the criteria for effective systems as put forth by Odden (2004) and Marzano (2014) but was stymied in the validity category. "There just have not been enough reported uses of the system or methods tracking its effectiveness at improving teacher instructional practices" offers Toby King, senior consultant for educator evaluation and support with the Colorado Department of Education (T. King, personal communication, March 13, 2013). For a system of teacher evaluation to be defensible,

either professionally or legally, it must be fair, i.e., the judgments made about a teacher's practice must accurately reflect the teacher's true level of performance and accurately predict student achievement (Danielson, 2012). The Colorado model for evaluating teachers fell short in terms of acknowledging and rewarding growth. A fear looming in many districts was the state's evaluation system would serve to acknowledge short-comings and be punitive in nature (Meyer, 2010). When decisions have high stakes consequences, such as teachers potentially losing their jobs, they heighten the demand for measures of teacher performance that are valid, reliable, credible, fair, and legally defensible (Stumbo & McWalters, 2011).

When the core purpose of evaluation is growth and improvement of teacher instructional practices, professional development tightly linked to performance standards and differentiated based on individual teacher needs is required. The criteria offered by Odden (2004) and Marzano (2014) represented in the Colorado model of teacher evaluation offered suggestions for system development that focused on practices that were clear, consistent, and likely to produce student achievement gains.

Standards-Based Performance Evaluation

The Colorado Model for Teacher Effectiveness (Educator Effectiveness Act, 2010) evaluation rubric intentionally included the Marzano characteristics for improvement-oriented evaluation systems in its system design (T. King, personal communication, March 3, 2013). Some districts in Colorado have intentionally included these elements in their local systems of teacher evaluation to reflect standards believed to capture instructional quality and also improve student learning through ongoing improvement of teacher practices (Colorado Legacy Foundation, 2013). The

comprehensive standards reflected in local systems might provide goals for instructional practice perceived as relevant and worth pursuing and provide a common vocabulary for discussing current notions of quality instruction (Heneman & Milanowski, 2003).

As a result of the *Race To The Top* federal grant competition, states were challenged to "design and implement rigorous, transparent, and fair evaluation systems for teachers and principals" (U. S. Department of Education, 2009, p. 34). Challenges faced by recipients of the *Race To The Top* grant were to strengthen the quality of evaluators in order to enhance the reliability of the evaluation and the objectivity of the results and another was to "articulate the full range of teacher practices and student outcomes that we want from our education system- and determine how we can measure them" (Stumbo & McWalters, 2011, p. 13).

Standards-based performance evaluations offered promise for comprehensive change as they could serve two main purposes: (a) assessing how effectively teachers were doing their jobs, and (b) providing valuable information to drive professional growth (Papay, 2012). The challenge was to merge the need to satisfy legitimate demands for quality assurance while promoting professional learning and improvement of instructional practices (Danielson, 2011). Districts desired an approach could can distinguish more and less effective teachers in a fair and reliable way; a method for using teaching standards to guide recruitment, selection, and promotion of teachers; and a strategy to enhance the core of a district's professional development system (Odden, 2004). Data derived from teacher evaluation could help identify and support teachers who are struggling and then could help evaluate them (Duncan, 2009). If teacher evaluation is meant to improve student learning systematically, evaluation must be used

as a tool to promote continued teacher development by assessing performance and identifying areas in which to improve instructional practice (Papay, 2012). A commitment to professional learning is important, not because teaching is of poor quality and needs to be fixed but rather because teaching is so hard it can always be improved (Danielson, 2011).

The theory of action or what is anticipated to occur when evaluation systems include teaching standards that describe effective instructional strategies is that teacher effectiveness will improve and student achievement will improve as well (Holtzapple, 2003; Marshall, 2005; Odden, 2004). The move toward sharpening the ability to evaluate teachers has gathered steam through the development of comprehensive rubrics of teaching practices (Bambrick-Santoyo, 2012). Teacher performance evaluation has the potential to improve instructional effectiveness and student learning by enabling teachers to receive high quality guidance and feedback on performance standards, thus improving their instruction (Donaldson, 2009).

Rubrics utilized in standards-based evaluation systems have several advantages over traditional evaluation instruments: (a) they force the evaluator to give the teacher clear feedback with respect to a standard, (b) they are more informative by telling teachers where they stand with a detailed description of what performance looks like at each level of proficiency, and (c) they may counteract grade inflation if it is clear very few teachers are at the advanced level (Marshall, 2005). Rubrics provide a method to conduct performance observations with integrity and skill by justifying all indicators with specific evidence (Heneman & Milanowski, 2003; Johnson & Fiarman, 2012).

Reliability of Performance-Based Measurement Tools

Before using any measurement and evaluation tool, it must be clear that the tool measures consistently, i.e., it is reliable, measures what it intends to measure, and is valid (Namaghi, 2010). When a teacher's evaluation is conducted using a tool that is invalid, evaluation cannot provide effective formative feedback (Namaghi, 2010). Evaluators need to know what effective teaching looks like and that teacher practices aligned with Common Core expectations might look differently from what has been experienced in the past (Varlas, 2012).

In 1996, Charlotte Danielson introduced her first version of the *Framework for Teaching*. Initially, the intent of the framework was to provide observation-based evaluation of first year teachers for the purpose of licensing (Danielson, 2013). The framework has since surfaced as a reliable method for observing classroom instructional practices demonstrated by teachers and improving the quality of professional conversations (Sporte, Stevens, Healey, Jiang, & Hart, 2013; Taylor & Tyler, 2012). Variations of the Danielson framework have become evident in district and state rubrics since empirical studies have shown each component of the *Framework for Teaching* was associated with improved student learning and a consistent definition of good teaching (Danielson, 2011, 2012). The *Framework for Teaching* identified those aspects of a teacher's responsibilities that have been documented through empirical studies and theoretical research as promoting improved student learning and sought to define what teachers should know and be able to do in the exercise of their profession (Danielson, 2013, p. 3). School districts and teachers rely on comprehensive frameworks based on irreducible elements of good teaching to guide their daily practice (Marshall, 2005;

Mielke & Frontier, 2012). The *Marzano of Teacher Evaluation Model* (2014) is another way to address the things teachers do in classrooms that have a direct impact on student achievement. It provides a different framework for establishing a common language across classrooms as well as being an instrument for observation and feedback (Marzano, 2014). With a clear and accepted definition of good teaching, teachers know how their performance will be evaluated and observers know what to look for (Danielson, 2012; Marzano, 2014). When comprehensive frameworks are used throughout the school year to collect data related to teaching, teachers can reflect on their practice and identify specific instructional strategies they can work on to improve their repertoire of skills (Mielke & Frontier, 2012).

Clear standards, highly-qualified and well-trained evaluators, and a focus on evidence can help remove subjective bias within evaluations (Papay, 2012). However, evaluators are often asked to do too much; problems with the implementation of evaluation systems were reported relating to the heavy workload they imposed on evaluators (Heneman & Milanowski, 2003; TNTP, 2013). Often there is too much to look for and think about in a short amount of time when rubrics are complex. Evaluators are expected to rate teachers on every aspect of performance the state or district wants to evaluate in addition to providing detailed feedback for teachers that can help address specific areas (TNTP, 2013). The time evaluators are required to devote to evaluations might also be a deterrent to accurate and high quality feedback as evaluators reported using four to six hours per formal observation including the components of pre- and post-conferencing, data management, and data analysis (Donahue, 2013; Sporte et al., 2013). Conflict related to inadequate communication between evaluators and teachers was also

reported as a stumbling block and cause for frustration within evaluation systems (Varlas, 2012).

Reliability of evaluation is increased when a comprehensive approach to observer training includes calibration and recalibration of observation and scoring techniques to enhance inter-rater agreement, increased attention to the interactive skills of professional conversation, and invitations for teachers to reflect on their practice and strengthen it in ways described by the instructional frameworks they use (Danielson, 2012). There was agreement among researchers that prospective evaluators need to achieve sufficient agreement during training using authentic classroom observations in order to be certified (Kane & Cantrell, 2012; Milanowski, 2004; Taylor & Tyler, 2012). However, issues with accuracy and leniency surrounding the practices demonstrated by evaluators were often reported (Milanowski, 2004; Taylor and Tyler, 2012). Danielson (2011) suggested that observers required multiple opportunities to practice using the observation rubric effectively and to calibrate their judgments with others.

Enhancing Self Reflection and Feedback

The Colorado Model for Teacher Effectiveness evaluation process began with a self-assessment by teachers on how well they perceived they were meeting the professional practices underscoring each element and quality standard of the Colorado model educator system for teachers (CDE, 2013). For many teachers, this was the first indication of potential avenues for professional growth during the school year and became the starting point for teacher-evaluator conversations. Heavy emphasis was placed on encouraging the evaluatee to be "thoughtful, comprehensive, and honest in the approach to self-assessment" (CDE, 2011, p. 18). This emphasis reflected the belief that

genuine development of effective instructional practices for teachers came about through self-awareness, reflection, and open-mindedness to other approaches and styles (Cosh, 1999). Although cognitive factors such as self-efficacy (Bandura, 1977) could affect how feedback is perceived and applied to instruction, the opportunity to self-reflect and engage in professional dialogue with peers helps teachers clarify how they should invest their efforts to grow in the profession (Kimball, 2002; Mielke & Frontier, 2012). Teachers might use an evaluation framework as the starting point of a comprehensive self-assessment process where specific skills for improvement are identified and feedback is obtained through such activities as peer observations, video analysis, peer discussion, and student surveys (Mielke & Frontier, 2012). Since professional learning is learning and learning requires the teacher to be an active participant in the process (Danielson, 2012), allowing teachers to generate data about their own areas of focus and establish their own improvement goals increases teacher motivation and engagement within their own practice (Mielke & Frontier, 2012). Cosh (1999) proposed that all performance-based evaluation systems should encourage self-reflection and self-awareness about teaching.

Heneman and Milanowski (2003) and Kimball (2002) noted promising findings in that teachers reported examples of changes in their instructional practices in response to classroom observations and the feedback they received. Recent findings from the measures for effective teaching project (Kane & Cantrell, 2012) suggested that evaluation practices should be fine-tuned to deliver the most usable feedback to teachers.

The idea that teacher evaluation is a driver of teacher development makes regular feedback essential. Quality of feedback is a central aspect of teacher evaluation that has

been shown to relate to perceptions of overall evaluation quality (Kimball, 2002). Stiggins and Duke (1988) identified attributes of feedback with a high correlation to teacher perceptions of evaluation quality including perceived evaluator credibility, quality of ideas, depth of information, persuasiveness of rationale for suggested changes, trustworthiness of evaluator, perceived relationship with the evaluator, and perceptions of evaluator capacity to demonstrate needed changes. When feedback is tied to evidence and agreed upon criteria, the opportunity for teacher reflection and improvement might be enhanced (Bambrick-Santoyo, 2012; Danielson & McGreal, 2000). According to Bambrick-Santoyo (2012), feedback providing corrections and improvement that build true talent cannot happen once every six months. It must loop relentlessly. Feedback and evaluation cannot change real classroom practices unless teachers build the skills needed to make such change. By identifying evidence of agreed upon instructional practices, evaluation enables evaluators to provide feedback and an opportunity for teachers to reflect on the use of past and future instructional practices (Mielke & Frontier, 2012). Taylor and Tyler (2012) found that micro-level evaluation feedback was more important to lasting performance improvements than the final overall teacher evaluation scores.

In a study of teacher perceptions of fairness in an evaluation system built upon the Danielson *Framework for Teaching*, Kimball (2002) developed a conceptual model that provided a view of how classroom observations might translate into improved teacher practices. Drawing upon earlier studies by McLaughlin and Pfeifer (1988) and Kelley (1999), Kimball addressed the idea that feedback alone within an evaluation system is not enough to promote instructional change. Kimball's model puts forth the idea that whether or not teachers use the feedback provided through observation and evaluation is

influenced by a common, comprehensive understanding of teaching, the quality of feedback, perceived fairness in the system, and enabling conditions in the organization. If evaluation is to be a vehicle to achieve the mission of public schooling (Phillips & Weingarten, 2013) and if teacher evaluations are a potentially powerful lever of school improvement (Toch & Rothman, 2008), use of Kimball's conceptual model might help one find collective meaning in the work of Stiggins and Duke (1988), Danielson and McGreal, (2000), Mielke and Frontier (2012), and Taylor and Tyler (2012) and use it to develop stronger, more effective evaluation systems.

Challenges Within Existing Evaluation Systems

Teacher evaluation has been criticized as superficial (Stiggins & Duke, 1988) or based on simplistic criteria with minimal relevance to what teachers need to do to enhance student learning (Danielson & McGreal, 2000). Evaluation systems have been discredited for lack of district- or school-level commitment, criteria based on narrow conceptions of teaching, inadequate feedback, and perceived subjectivity (Kimball, 2002); they have been labeled problematic due to lack of shared values and assumptions about good teaching (Danielson & McGreal, 2000), low validity (Medley & Coker, 1987), an emphasis on following procedures rather than improving performance (Johnson, 1990), limited feedback (Danielson & McGreal, 2000), low utility to teachers as a means for improving performance (Johnson, 1990), and pervasive mistrust on the part of teachers toward evaluation (Johnson, 1990; Peterson, 2000). Toch and Rothman, in their 2008 report *Rush To Judgment*, used the term "capricious" to describe the common practice of drive-by evaluation (p. 2). Problems reported for evaluators included a heavy workload, lack of clear guidelines on the content, quantity and quality

of portfolios and other evidence, and uncertainties on the part of teachers and evaluators about the interpretation of quantifiers like "consistently" and "frequently" in the rubric language (Heneman & Milanowski, 2003). If teacher evaluation holds the key to improving instruction and consequently growing student achievement as some research leads us to believe, each of these limitations must be addressed and refined.

Conclusion

Results as far back as Wright et al. (1997) implied a need to improve education through improving the effectiveness of teachers, more than any other single factor, and high quality classroom observations pointed to potential value for improving instruction (Taylor & Tyler, 2012). If an evaluation system is operating effectively, one should find, over time, that professional development opportunities are available for teachers to learn new practices; teachers' practices are conforming more to teaching standards; and as instruction evolves toward what is described in standards, student performance increases on a continuous basis (Odden, 2004). But challenges within existing systems continue to surface.

With all of the research that supports a focus on improving teaching practices in order to improve student outcomes, research identified little connection between evaluation and the improvement of student learning outcomes (Chingos & Peterson, 2011; Long, 2000). Chingos and Peterson (2011) stated, "In general we find it much easier to pick a good teacher than to train teachers to make them more effective" (p. 464). The 2009 RAND (Buddin & Zamarro, 2009) study findings suggested that traditional measures of teacher quality did not predict classroom performance.

Although research is inconsistent in terms of viewing teacher evaluation systems as sources of information on teacher instructional behaviors that affect student learning (Milanowski, 2004), mounting evidence indicated that rigorous standards-based evaluations could improve teacher effectiveness and rigorous evaluation systems might boost teacher effectiveness and student achievement (Papay, 2012). Odden (2004) reported top-rated teachers produced the most learning gains and strong linkages between teacher evaluation scores and student learning gains were found. On average, Odden found that teachers with higher evaluation scores produced more student learning gains than predicted based on prior test scores and demographics of students in their classrooms. Taylor and Tyler (2012) found that evaluation scores predicted student achievement. The Eagle County School district in Colorado reported similar predictive findings in their informal comparison of teacher evaluation scores and student Colorado Student Assessment Program (CSAP) performance (Smyser & Wodlinger, 2013). Perhaps the most obvious benefits of teacher evaluation systems were the development of a consistent definition of good teaching, opportunities to engage in meaningful conversations about practice, and a focus on what matters within classroom instruction (Bambrick-Santoyo, 2002; Danielson, 2011, 2012; Heneman & Milanowski, 2003; Kimball, 2002; Marshall, 2005, 2012; Mielke & Frontier, 2012; Odden, 2004).

Quantitative research is beginning to support the mantra that teachers are the most important school-level factor in promoting student achievement (Papay, 2012) but very little is known about how information regarding teaching practices or the experience of being evaluated might change teacher effort and effectiveness (Taylor & Tyler, 2012). If our educational beliefs about students reflect thinking that every child can learn and

every school can succeed, we must provide evaluation systems teachers find meaningful and from which they can learn (Duncan, 2009). We must use processes that are rigorous, valid, and reliable, and also engage teachers in activities that promote learning, e.g., self-assessment, reflection on practice, and professional conversation (Danielson, 2011). Developing a collective understanding of effective teaching and turning the focus on how teaching could be improved by evaluation might enable districts and states to build systems including qualities that replicate best evaluative practices.

CHAPTER III

METHODOLOGY

Policymakers and practitioners alike need to better understand how teacher evaluation informs change in teacher practice and whether it has the potential to lead to more effective instruction (Taylor & Tyler, 2012). Very little is known about how the experience of being evaluated might change teacher effort and effectiveness (Taylor & Tyler, 2012). This study focused on how the components and processes of evaluation systems were perceived by teachers to change their instructional practices.

The purpose of this case study was to obtain qualitative data surrounding teacher perceptions of how their instructional practices might be impacted by an existing teacher evaluation process. An objective of this research was to capture the teacher voice, which would enable system designers as well as evaluators to respond to ways teachers felt the system of observation and evaluation impacted their professional growth. This chapter clarifies the design of the study, sample selection, data collection, and methods of data analysis. The following research questions guided this investigation:

Q1 How do novice teachers feel the district evaluation system impacts their classroom practice?

Q2 How do experienced teachers feel the district evaluation system impacts their classroom practice?

Design of the Study

Creswell (2007) described qualitative research as an approach to inquiry "that begins with assumptions, worldviews, and a theoretical lens through which to study the meaning individuals or groups assign to a social or human problem" (p. 37). Merriam (2009) described qualitative study as "research focused on discovery, insight, and understanding from the perspective of those being studied and offering the greatest promise of making a difference in people's lives" (p. 1). Morse, Barrett, Mayan, Olson, and Spiers (2002) suggested that qualitative research is "iterative rather than linear so that a good qualitative researcher moves back and forth between design and implementation to ensure congruence among question formulation, literature, recruitment, data collection strategies, and analysis" (p. 10). Since qualitative research focuses on process, meaning, and understanding, it was the most appropriate design for this study. Shields (2007) offered that qualitative studies do not attempt to discount what cannot be discounted or simplify what cannot be simplified. The idea of *not* simplifying teacher perceptions of the system in which they are evaluated was relevant to this study since the way in which evaluation systems might be perceived is complex and based on one's personal experience.

From an epistemological perspective, this study was viewed through a constructivist lens. Social constructivism--the seeking by an individual of an understanding of the world in which they live and work (Creswell, 2007)--guided this research toward an understanding of the subjective meanings of teacher evaluation. Emphasizing the importance of the participant's view and highlighting the meaning held personally by the individuals in the study were central to the study's purpose (Creswell,

2008). The constructivist approach illuminated how teachers perceive an existing system of evaluation's effect on improving their teaching and brought to light perspectives held by teachers who find tremendous value or little value in the existing evaluation system as a means to improve their instructional practice.

Case study design was used in this research to examine perceptions of teacher evaluation within one Colorado school district--a bounded system based on time, location, and the district's history of using teacher evaluation as a means to improving instructional practices. Case study, an in-depth exploration of a bounded system based on extensive data collection (Creswell, 2007), revealed previously undisclosed perceptions about the phenomenon of teacher evaluation, "knowledge to which we would otherwise not have access" (Merriam, 2009, p. 46). Case study was the appropriate method for this research as it provided thick description and grounding, holistic, and simplified data to be considered by the reader (Guba & Lincoln, 1981). Case study was further selected for this research because although the researcher sought to understand the perspectives held by individual teachers, the similarities and differences between each individual experience within the collection of experiences provided the most compelling information (Stake, 2006). Through the study of each individual experience, perceptions were uncovered that were rooted in specific context. The perceptions held by individuals became more developed through comparison within the case and through reader interpretation. Case study promoted investigation of the phenomenon of teacher evaluation within its real-life context when the boundaries between context and phenomenon are not clear (Yin, 2009). Ethnographic or narrative research design, while likely to illuminate the story behind the evaluation experience, would not have provided

data specific to what worked within the evaluation systems and how teachers perceived the system to work for themselves as individuals.

Although case study was selected with intention and purpose, existing limitations were carefully considered. In case study, results might be limited to describing a phenomenon rather than predicting future behavior (Yin, 2009). A common concern is case studies provide "little basis for scientific generalization as they may be generalized to theoretical propositions but not to populations or universes" (Yin, 2009, p. 15). Hamel, Dufour, and Fortin (1993) made the point that "case study has been faulted for its lack of representativeness and its lack of rigor in collection, construction, and analysis of empirical materials and suggested that it may be linked to problems of bias introduced by the subjectivity of the researcher" (p. 23). Case study research might be limited further by the sometimes too lengthy, too detailed, or too involved final product for policymakers or practitioners to read and use (Merriam, 2009). When the term "case study" is not used precisely, it becomes a catch-all category for studies that are clearly not experimental, survey, or historical (Merriam, 2009, p. 54).

Strengths, however, outweighed limitations in terms of case study. The case study results in a rich and holistic account of a phenomenon. Case study offers insights and sheds light on meaning that can be construed as tentative hypotheses and can expand a reader's experiences to help structure future research (Merriam, 2009, p. 51). Readers can learn vicariously from an encounter with a case through the researcher's narrative description; it becomes the reader, not the researcher, who determines what can apply to his or her context (Stake, 2005, p. 455).

The case examined in this study was one Colorado school district; participants were selected from within its ranks. Mountain School District[1] (MSD) has been selected since it is an example of a "process, issue, or concern" within existing evaluation systems (Stake, 2006). Mountain School District, a school district of over 15 schools, offered a wide pool of teachers for participant selection who have experienced the existing teacher evaluation system in multiple ways and from multiple perspectives. Mountain School District possesses a lengthy history of using an evaluation system in an effort to promote effective teaching and raise student achievement. A brief description of the evaluation system utilized in MSD is provided in Appendix A. Institutional Review Board approval is presented in Appendix B.

Pilot Study

The researcher's 2013 pilot study examined responses to questions similar to the research questions posed in this study and sought to compare teacher perspectives of evaluation experiences within MSD and the teacher's previous district. A much smaller sample consisting of one teacher from each school model in the district (elementary, middle school, high school, and a K-8 school) was used. The pilot study sampled teachers new to MSD but not new to teaching. The researcher looked for patterns and themes to potentially influence the redesign of an existing evaluation system. Changes to interview questions, methods of recording interviews, use of multiple sources of evidence, and selection of participants were influenced by the pilot study. The pilot study informed the design of the current study in that it helped to refine data collection plans with respect to the content of the data and the procedures followed. The pilot study

[1] Mountain School District is a pseudonym.

informed the development of relevant lines of questioning, illuminated prevailing perspectives, and initiated new learning for the researcher in terms of sample selection, data collection, and field procedures. The findings of the pilot study uncovered new pathways for the interview protocol and clarified propositions that underlie this study but did not warrant generalization within the district. A common theme that emerged from the pilot study interviews was a perception that feedback existed as a strength within the evaluation system of MSD and led to changes in classroom practices. Another theme gleaned from the study was the participant perspective that feedback must be informed by set criteria, be specific, be constructive, include suggestions for change or growth, occur with frequency, and include follow up (Donahue, 2013).

Criteria related to the establishment of themes within the pilot study required perceptions to be evidenced at least three times. The criteria for the establishment of themes prevented topics such as the need for feedback to come from trusted sources, visiting colleagues' classrooms as a method for giving and getting feedback, utilizing specific criteria outlined in the professional practices rubric to plan effective feedback, experiencing multiple evaluators at multiple times of the year, and having an opportunity for both written and verbal feedback from being topics of specific focus. Many of these topics resurfaced within the current study.

Sample Selection

Mountain School District, the case under study, gained notoriety for its comprehensive evaluation system. It was an ideal case for this study as it offered a significant contribution to knowledge and theory building surrounding teacher evaluation (Yin, 2009, p. 47). The individuals and sites presented in this case purposefully informed

an understanding of the research problem (Creswell, 2007). Within the case, stratified random selection was used to select participants and garner the perspectives of a wide range of teachers across grade levels, years of experience, and areas of content instruction. In stratified random sampling, the researcher stratifies or divides the population according to a specific characteristic and then, using simple random sampling, samples from each subgroup (Creswell, 2007). Mountain School District employs approximately 245 secondary teachers and approximately 200 elementary teachers. Simple random sampling without stratification might yield a biased view or perspective from teachers in high schools or middle schools. To avoid this limitation, stratification was used to ensure that the stratum desired was represented in the sample in proportion to the existence in the population (Creswell, 2007). Since individuals selected through random sampling procedures gave every individual an equal chance of being selected, the bias among the population was equally distributed among the people chosen (Creswell, 2007). To represent the complexity of the case, it was important to use stratified random sampling as it allowed for differences in school level, experience, or years of teaching, high and low scoring teachers, content area of instruction, and roles within the school system. Some teachers were also evaluators and stratified random sampling gave those teachers an equal chance of being randomly selected from all of the teachers in their buildings. Random selection among the participant sample was warranted as the study sought to examine the perspectives of teachers across a wide range of contextual and evaluation experiences.

Within MSD, two teachers were randomly selected from each school. Stratifying the school district by individual schools was necessary to uncover trends that might be

influenced by school-based attitudes toward teacher evaluation or by individual evaluators. Random selection from within each school provided an equal opportunity for teachers who have had positive or negative experiences with evaluation, received helpful versus unhelpful feedback, have deep understanding of the evaluation system versus limited understanding, have varying levels of relationships with evaluators, or have evaluators with varying levels of expertise.

Participants were selected randomly and placed into an ordered list using a Kagan random name selector tool (Kagan & Kagan, 2009). The researcher contacted each participant, invited them to participate, and provided a description of the study. In the event one or both teachers selected from a school did not wish to participate, the researcher invited the next potential participant from the randomly generated list. The identity of the participant was known to the researcher only unless the participant wished to share his or her experience with colleagues. Preserving confidentiality to the greatest extent possible resulted in more candid conversations with teachers and avoided lack of responsiveness--the greatest hidden threat to validity (Morse et al., 2002). Within MSD, one school (a charter school) was not required to adhere to the district evaluation protocol; therefore, no participants were invited from that school. Additionally, participants were not included from the school at which the researcher is the principal. The final sample used for the study included two teachers from 15 district schools. An example of the consent form signed by participants is provided in Appendix C. A complete list of participant pseudonyms aligned with participant school level, instructional content area, and classification as “novice,” “new-to-MSD,” or “experienced” is presented in Chapter IV in Table 1.

Limitations of stratified random sampling might include not having enough teachers willing to participate in the study. The researcher made repeated attempts to engage two participants from each school. Ultimately, 39 individuals were extended an invitation in order to obtain commitments from 30 individuals who expressed an interest in participating in this study. The final sample included representation from most grades and levels. Kindergarten through grade 12 teachers, special educators, physical education teachers, art teachers, language specialists, elementary teachers, secondary teachers, and teachers of literacy, math, science, and social studies were represented.

Data Collection

Case study relies on multiple sources of evidence and data converging in a triangulating fashion (Yin, 2009, p. 18). Semi-structured interviews were the primary method of data collection. Additional and related documents such as self-reflection forms and pre- and post-conference documents were anticipated to be introduced by participants but were not evidenced. Maintenance of an audit trail, a chronological record providing evidence of the sequence of events and data collection within the study, and development of a researcher's journal, a narrative explaining how decisions within the research are made, informed the collection of evidence and supported the organization of data within the case study database. Triangulation between multiple data sources, the comparison and cross-checking of data collected through interviews with people holding different perspectives, helped address potential problems of validity because multiple sources of evidence essentially provided multiple perspectives of the same phenomenon (Merriam, 2009; Yin, 2009).

Role of the Researcher

The researcher, as the primary instrument of data collection and analysis within case study design, relied on her instincts and abilities throughout the research effort (Merriam, 2009, p. 52). This included an effort on the part of the researcher to look for data that supported alternative explanations (Patton, 2002). The researcher adhered to guidelines of ethical inquiry that did not disrupt or manipulate the normal life experiences of participants, incorporate any form of intrusive behavior, or involve deception. The researcher documented progress throughout the study by maintaining a detailed account of data collection within an audit trail and by maintaining a researcher's journal. Individual responses were gathered through the researcher's commitment to open and positive relationships before, during, and after the one-on-one interviews.

Interviews

Interviews, unlike surveys, are guided conversations rather than structured probes. In case study, the actual stream of questions is likely to be fluid rather than rigid (Rubin & Rubin, 1995). Semi-structured interviews were conducted because they allowed for modifications to be made throughout the study in an as-needed and ongoing manner. While questions followed a line of inquiry from the case study protocol, the semi-structured interview allowed the researcher to modify questions if relevant information based on the study's initial propositions and research questions was not forthcoming. The interview was an essential source of evidence for the case study because this study was about human perspectives.

Each interview was conducted face-to-face and lasted 60-90 minutes. Each interview took place at a public building of the participant's preference. Many

participants selected libraries, likely due to the low level of noise and the greater opportunity to use secure rooms. However, some participants invited the researcher to their homes and still others requested the interview take place at the school in which they worked. Participants were informed that the interview would be audio-taped using the researcher's iPad and iPhone to ensure a backup for capturing the conversation. Once each interview was securely recorded and transcribed by the researcher, it was transferred to an external hard drive, transmitted to an external transcriber via Dropbox, and deleted from both recording devices. Transcripts of each interview were made available to participants for verification of accuracy. Member checks, the solicitation of feedback on emerging findings from the people interviewed, ruled out the possibility the researcher misinterpreted the meaning of what participants said (Merriam, 2009).

Questions asked within this study related to the experiences teachers have had with the evaluation system in MSD including process elements such as pre-conferences, post-conferences, coaching, and scoring. Questions that addressed how teachers perceived the evaluation system to influence changes in their instructional practices comprised the bulk of the interview protocol.

Although interviews provide insightful information, they could be limited by poorly articulated questions, response bias, and reflexivity (the interviewee gives what the interviewer wants to hear). Interviews might be further limited if participants feel their privacy has been invaded, they are embarrassed by certain questions, and they tell things they never intended to reveal (Stake, 2005). Recognizing these limitations, the researcher maintained a belief in the value of interviews to glean the multiple and varied perspectives that brought value to this research.

Since participants in this study were selected through stratified random sampling, it was necessary to gather background information to determine how well the population of the case was represented. Context building questions (see Appendix D) served to gather information that described each participant and was helpful in interpreting responses. Other questions directly related to the purpose of this study followed the initial phase of context building. These questions served as a springboard for discussion on a variety of evaluation related topics (see Appendix E). The interview closed with a final question asking participants to assign a rating from 1-10 on the perceived impact the existing system had on their instructional practices. Additional assurances were provided that although the evidence gleaned from this study might be used to inform change in the existing evaluation system, comments would not be attributed to the participants from whom the comments were derived. Interview was selected as the primary method of data collection because the ensuing discussion illuminated feelings and perspectives that might not have been experienced personally by the researcher or future readers.

Data Analysis

Analysis of data collected throughout this study addressed the most significant aspects of the case as well as rival interpretations and contrasting perspectives of individuals within the case. The questions asked during the interviews were attempts to collect evidence from multiple perspectives on the topic of teacher evaluation. Data analysis followed the original propositions leading to the case study.

Analysis began with open coding and identification of comments used to inform the construction of themes. Open codes grouped together and based on the researcher's interpretation and reflection on meaning sparked the development of themes. The

researcher returned to the data to align evidence around identified themes, not for generalizing beyond the case but for understanding the complexity of teacher evaluation within the case study. Using a model of coding suggested by Creswell (2007), data were inductively analyzed beginning with reading through the text data, dividing the text into segments of information, labeling the information with codes, reducing overlap and redundancy of codes, and collapsing the codes into themes (Creswell, 2008, p. 251).

Since all data analysis is ultimately inductive and comparative (Merriam, 2009), data were further analyzed using within-case analysis suggested by Yin (2009). This technique treated the evidence surrounding each individual as a separate study and included within-case, cross-examination of topics or themes as they evidenced themselves.

Trustworthiness

The trustworthiness of qualitative research is often questioned by positivists because the concepts of validity and reliability are not addressed in the same ways they are addressed in quantitative research (Shenton, 2004). In qualitative research, Lincoln and Guba (1985) substituted reliability and validity with the parallel concept of trustworthiness containing the following four aspects: credibility, dependability, confirmability, and transferability. Merriam (2009) suggested focusing on strategies to establish trustworthiness throughout the study rather than at the end when it might be too late to correct errors that influenced the study's trustworthiness.

Guba and Lincoln (1981) proposed the enhancement of credibility through the establishment of structural coherence, which ensures there are no unexplained inconsistencies between the data and their interpretations. Eisner (1991) used the term

structural corroboration to describe the technique used by researchers to synthesize multiple types of data in order to support or contradict an interpretation. Krefting (1991) offered the idea that credibility might be increased when contradictions, rival explanations, or deviant cases can be accounted for and enhanced through the interview process. In this study, triangulation between interview data gleaned from a wide variety of participants, evaluation-related documents, and member checks enhanced the credibility of the study. Member checking, a method of providing a written transcript of the recorded interview for each participant to review and provide feedback on its accuracy, was used as the primary method for ensuring the researcher accurately represented the perceptions shared during the interview. Individual viewpoints and experiences could be verified against others; ultimately, a rich picture of the attitudes, needs, or behaviors of those under scrutiny was constructed based on the contributions of a range of people (Shenton, 2004). Site triangulation was achieved by the participation of teachers within most schools in MSD to reduce the effect of particular school-based factors. When similar results emerge at different sites, findings have greater credibility in the eyes of the reader (Shenton, 2004). Credibility was also enhanced by stratified random sampling, which attempted to ensure that all parts of the population were represented in the sample.

Lincoln and Guba (1985) suggested that dependability might be confirmed through a single audit, compelling the researcher to prepare a detailed audit trail. An audit trail was kept throughout the duration of the research and was used to document the course of development of the completed analysis. The audit trail included a detailed account of how the study's method was carried out in the field while the researcher's

journal chronicled how and why decisions regarding the course of the study were made. The database used to enhance reliability was directly informed by the dependability of the audit trail and researcher's journal and provided a detailed account of how evidence became aligned to developing themes.

Confirmability addresses the reduction of the impact of researcher bias. Integrating the evidence from multiple data sources and sites, maintaining an audit trail and researcher's journal, as well as maintaining a database of evidence and developing themes confirmed the integrity of the study and findings. The researcher's admitted bias toward the potential of evaluation systems to improve instruction was carefully and completely contained to avoid contaminating the study's findings.

In terms of transferability, it was not the intent of this study to confirm generalizations regarding teacher experiences with evaluation systems. The purpose of this study was to analyze data regarding perceptions held by teachers in the MSD district about how they felt the evaluation system influenced changes in their instructional practices. It would be the responsibility of other readers to determine whether the results of this study were relevant and of interest within their own context. They might wish to conduct additional studies to determine whether or not the findings from this study could be transferred across other districts or systems. The audit trail, researchers journal, and database showing the collection and alignment of evidence with themes might allow readers to determine for themselves if these findings could be transferred across a variety of contexts.

Researcher Bias and Assumptions

The purpose of this study was to examine teacher perceptions of how evaluation systems might inform change in teacher instructional practices. The researcher is a principal at one of the elementary schools in MSD. She has held a variety of other positions in the school district such as mentor teacher, master teacher, and district instructional coach. She has served on multiple committees whose goals centered on the evolution of the district evaluation process. The researcher has a keen interest in examining how an existing system, and the aspects of observation and evaluation embedded within it, actually impact teacher instruction and student growth. It is fair to say that the researcher carries a bias toward the perceived value of teacher evaluation systems to improve teacher practice and student achievement when designed and used effectively.

A plethora of research exists laden with detailed information about the value of evaluation systems (Bambrick-Santoyo, 2012; Danielson, 2011, 2012; Heneman & Milanowski, 2003; Kimball, 2002; Marshall, 2005, 2012; Mielke & Frontier, 2012; Odden, 2004) but the research questions in this study did not specifically target how teachers perceived the systems they were bound within to influence their personal and individual choice of classroom practices. The researcher hoped this study would elucidate themes and topics that might lead to increased effectiveness of an existing evaluation system. The researcher recognized that as a school district administrator, a member of various committees working on system redesign, and former professional developer, differences in experiences and opinions regarding the purpose and value of teacher evaluation existed. It required careful planning by the researcher to ensure this

study would not influence job security or other negative outcomes. Participants were provided assurance that their responses, which could potentially inform decisions made regarding system redesign, would not be attributed to them as individuals.

This study of teachers' perspectives regarding the impact of an existing evaluation system on individual classroom practices might clarify existing challenges in the MSD evaluation system. Findings might be transferable to other school districts and systems as previously unknown relationships might emerge from this case study and lead to a rethinking of teacher evaluation systems on a grand scale. As Colorado continues to revise the educator evaluation system to meet the criteria of the Educator Effectiveness Act (2010), the qualitative data gleaned from this study might provide authentic teacher perspectives regarding what works in evaluation systems. If one of the goals of the Educator Effectiveness Act is to provide a system that engages teachers in a model of continuous improvement, the findings from this study could be relevant to the design or redesign of systems intended to impact teachers and instructional improvement statewide.

CHAPTER IV

RESULTS

This chapter describes the results gleaned from this study of teacher perceptions regarding the impact of an existing evaluation on changing classroom instructional practices. The study, conducted within the Mountain School District of Colorado, included 15 of the 17 schools in the district. Only 15 schools were selected as the researcher is currently the principal at one of the remaining schools and one other district school, a charter school, was not required to adhere to the existing system of teacher evaluation. Using the procedure detailed in Chapter III, teachers were approached and invited to participate in this study. Overall, the researcher approached 39 teachers to obtain an interview with the desired number of participants. Those who rejected the invitation expressed being overwhelmed and having little time for additional commitments or expressed a lack of interest in the topic. One teacher made arrangements to participate on two separate occasions and showed up for neither. The majority of teachers contacted who did not want to participate merely did not respond.

Stratified random sampling was the method used to initially identify possible teacher participants; it ultimately yielded 25 experienced teachers and five novice teachers who agreed to participate. As data gathered from the study were analyzed, it became clear that three groups of teachers with differing perspectives existed rather than the two groups the research initially intended to examine. While there were clear criteria

delineating novice teachers from experienced teachers, a third category, new-to-MSD teachers, emerged as a group with their own set of perspectives. This group, comprised of seven teachers with four or more years' experience in the teaching profession outside of MSD, offered a comparative look between evaluation systems across multiple districts and states. Many teachers in this group expressed dismay as their years of experience did not match the scores they anticipated within the MSD system of teacher evaluation. It is important to recognize how these teachers veteran to the profession experienced a teacher evaluation system in a new district. Holly captured some of the surprise felt by new-to-MSD teachers when she had her first observation. She revealed,

> I was highly affected when I got my scores. And I was thinking like, you know…I've always gotten exemplary scores in every school that I've been at and when I got my scores I was in tears. I was like, "I can't sign this, I don't know what you're talking about. I've been doing this for a long time and this does not seem representative of me." The rationale for scoring was that '3' was professional and that did not register with me because I strived to do a certain job everyday with the kids that I affect and impact and a '3,' in the middle of the rubric, is kind of telling me that I'm a 'C.' Now, I have learned that's not the lingo of the district.

Mountain School District employs approximately 245 secondary teachers and 200 elementary teachers. To ensure a balance of perspectives, two teachers were selected from each school resulting in 14 secondary teachers and 14 elementary teachers. Of the teachers invited to participate from the K-8 school, the final sample included one elementary teacher and one middle school teacher. Approximately 15% of teachers in the school district are members of their school's instructional leadership team comprised of masters, mentors, and principals. The sample reflected the district-wide percentage of master and mentor teachers and included six teachers who are currently or have been master or mentor teachers (see Table 1). Within the 15 participating schools, two

teachers from each school were interviewed and provided data used to examine the following questions:

Q1 How do novice teachers perceive an existing evaluation system to impact change in their classroom instructional practices?

Q2 How do experienced teachers perceive an existing evaluation system to impact change in their classroom instructional practices?

The researcher further disaggregated years of experience to understand the perspectives of teachers new to teaching compared to being new-to-MSD and included the question: How do experienced teachers, new to MSD, perceive an existing evaluation system to impact change in their classroom instructional practices?

Table 1

Participant Descriptions

Pseudonym	Level	Content	Past or Present Master Teacher	Past or Present Mentor Teacher	Participant Classification
Farrah	Elementary	Reading Intervention	Yes	Yes	Experienced
Miranda	Elementary	Special Education	No	No	Experienced
Frida	Secondary	Special Education	No	No	Experienced
Michaela	Secondary	Math	No	Yes	Experienced
Rose	Secondary	Science	No	No	New-to-MSD
Nancy	Secondary	English Second Language	No	No	Novice
Anne	Elementary	Social Studies	No	No	Novice
Jane	Elementary	English Second Language	No	Yes	Novice
Casey	Elementary	Physical Education	No	No	Experienced
Brooke	Elementary	Grade 4	No	No	Experienced
Katrina	Secondary	Science	Yes	Yes	Experienced
Helen	Secondary	English	Yes	Yes	Experienced
Sally	Secondary	Grade 6	No	Yes	Experienced
Lucy	Secondary	Special Education	Yes	Yes	Experienced
Sage	Elementary	Art	No	No	Experienced
Kristine	Elementary	Grade 2	No	No	New-to-MSD
Kyle	Secondary	Grade 7	No	Yes	Experienced
Susan	Secondary	Special Education	No	No	New-to-MSD
Valerie	Elementary	Grade 1	No	No	New-to-MSD
Tess	Elementary	Grade 1	Yes	Yes	Experienced
Ginger	Elementary	Grade 3	No	No	Novice
William	Elementary	Grade 4	No	No	New-to-MSD
Sheryl	Secondary	English	No	Yes	Experienced
Mary	Secondary	Social Studies	No	No	New-to-MSD
Kate	Elementary	Kindergarten	No	No	Experienced
Taylor	Elementary	Grade 5	No	No	Novice
Joanne	Elementary	Physical Education	No	Yes	Experienced
Holly	Elementary	Grade 2	No	No	Experienced
Xavier	Secondary	Science	Yes	Yes	Experienced
Hannah	Secondary	English	No	No	New-to-MSD

Emerging Themes

The researcher interviewed teachers regarding their thoughts about what improved their instructional practice as well as how they perceived the existing evaluation system to support their professional growth. As interviews were transcribed and analyzed, the following five themes emerged related to the purpose of the study:

1. Effect of feedback
2. Quality of the relationships
3. Effect of the Professional Practices Rubric (PPR)
4. Effect of modeling
5. Effect of personal integrity and self-reflection.

The existing evaluation system in MSD includes an evaluation rubric as well as a system for providing teachers with coaching, modeling, and feedback. The system, although named an *evaluation* system, contains elements designed to support teachers and help them improve their practices in order to improve student achievement. Farrah, an experienced elementary teacher and member of her school's instructional leadership team, explained her perception of the long-term benefit of the existing system: "The rubric and the coaching improve teachers. Mountain School District makes great teachers who are highly employable anywhere they go." Casey, another experienced elementary teacher, offered a long range perception of the evaluation system:

> I think the system was implemented ideally to kind of weed out teachers that weren't performing or at least discover them so that they could be given additional training. However, what happened in those first few years was that experienced, good teachers left the district upset. Now I think it helps even experienced teachers become better. I know I've gleaned new things every year. There's always some refinement.

Teachers, both novice and experienced as well as experienced teachers new-to-MSD, overwhelmingly identified feedback and coaching using the PPR as the most impactful elements in terms of changing and improving their classroom practices. Of the 30 teachers in the study, 24 teachers discussed the ways in which the PPR provided them with clear expectations for performance. Participants also shared thoughts on how the relationship they had with their coaches (mentor or master teachers) provided them with a forum in which they were given feedback that enabled them to self-reflect and make instructional improvements. Self-reflection was identified by 80% of the novice teachers as the element bringing about the most change in their practice. However, in the larger group of experienced teachers, self-reflection, modeling, and coaching emerged from 18 out of 25 interviews as elements bringing about the most change. Jane, a novice elementary teacher, reflected on how she perceived the evaluation system to be effective at growing teacher instructional practices when the system was implemented the way it was intended:

> I think if all the components were implemented in the way that they were supposed to be, they would all be equally as effective in evaluation. But I think that the intent isn't always how it's played out, like how I had my three master observations in one week and it's supposed to be over months. And I love my principal but last year the first time she was in my room was in April and so she had her two drive-bys and then my evaluation and so I pretty much knew it was coming in that month and it's just scheduling and timing all of that jazz but I feel like if it was all implemented the way it was supposed to be and kind of spaced out it would be very beneficial. But when I have three master observations in one week and it's at the very, very end of the year, I don't have time to implement change and show that I was reflecting and trying to implement the change that we'd talked about in our post conference. So I think there is a little disconnect in the intent and the actual practice.

Perceptions that elements within the existing system served to inhibit growth instead of foster growth were prevalent and were shared within the context of the larger themes.

The most common complaint was the system did not take into account the specific elements within the job of a special education teacher. Other common criticisms were the system did not measure things teachers perceived as important such as humor or teacher/student relationships. Instructional risk-taking was mentioned by teachers as something they did not do for fear they would be evaluated on "that day." Casey, an experienced teacher, shared conversations she has had with teachers who say, "Wow, I would like to try this but I do not want to try something new knowing that it might be happening when I am getting evaluated." Michaela offered her personal experience: "I wanted to see how one of these lessons measured up on our evaluation tool. And so I was putting the lesson to a test but it backfired on me. I got low evaluation scores. I cannot say that I have taken that kind of risk again."

A final question asked all participants to provide data for a comparative look at the perceived usefulness of the existing system by teachers (see Figure 1). Teachers were asked on a scale of 1—*very little* to 10—*a lot* how much the existing system of teacher evaluation impacted their practices. All of the novice teachers in the study scored the system a 7 or above based on the perception the PPR provided them with a guide to expectations for good teaching within MSD. Experienced teachers consistently provided a ranking above a 5; those experienced teachers ranking the system with a 7 or higher revealed in a variety of words that "it impacts everything I do" and "it provides me with a framework on which to hang my instruction." For participants who ranked the system within the range of a 3 to 6.5, the reasons varied but a common perception was the system no longer impacted practice because the practices "have become part of everyday practice" or because "it is not the evaluation that impacts practice, rather it is the

professional development that is a result of the evaluation." Teachers new to MSD held perspectives that aligned with both novice teachers as well as experienced teachers. Of these seven teachers, five responded that the system impacted their professional practices "because the PPR provides them with feedback on best practices and guidelines for how to teach." One participant shared she found the system to be motivating, making her want to achieve high scores in all areas of the evaluation. Another noted, "It is so good to talk about instruction. I feel very fortunate for our system. It seems very structured, very purposeful. There is always a purpose and it is always about making us better teachers." However, one teacher found the system to have little impact on her professional growth and rated the system a 2 in terms of impact. The teacher offered, "It is not like I am working just because there is an evaluation." Data gathered over the course of more than 30 hours of interviews were positively and negatively reflected on the existing system and provided insight for future system modification and growth.

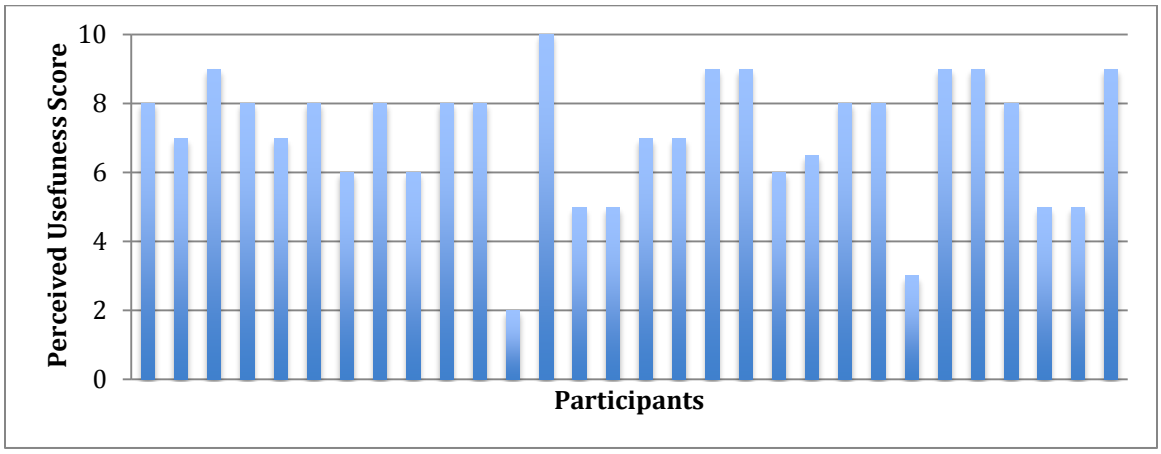

Figure 1. Perceived usefulness of existing system.

In the following section, the researcher used pseudonyms to present participant responses as they aligned with the emerging themes.

Effect of Feedback

Feedback was perceived by teachers to directly influence changes in their instructional practices but it was not perceived positively in all situations. The amount of feedback, the timeliness of feedback, the quality of the feedback, as well as the level of expertise possessed by the person offering the feedback were perceived by teachers to have an effect on whether feedback would be implemented.

Novice teachers in the study mentioned receiving feedback "whenever they ask for it," "asking specifically for feedback," and "trusting that my mentor teacher will have feedback to offer that is based on experience." Ginger, a novice teacher, shared that she "uses post conference feedback to reflect on how to make changes that are actually effective." A trend noted in responses from novice teachers was the desire for feedback that would "actually" help them improve. All of the novice teachers mentioned the type of feedback they received as integral to their growth. Nancy, a novice teacher, shared:

> I can trust that my mentor will have feedback based on experience to share with me. Generally I can get feedback whenever I ask for it. I was hoping for more feedback where I thought I needed it and I got more feedback on where I was strong. I'd like there to be less focus on what I am doing well and more focus on what I need to work on.

Anne, also a novice teacher, shared that while her experience with her mentor teacher did not provide her with the feedback she needed, she was able to ask her principal for coaching support and reported being pleased with the outcome:

> My mentor, being new to the position, wasn't probably really sure of what I needed and didn't know me at all or what I needed. I did ask specifically for feedback and for some more coaching rather than just evaluation. I asked specifically for that and I think it just worked out. My principal said from one of

> his ‘pop ins’ that he’d like to see this and then he gave me a link with some suggestions and strategies and it just so happened the next time he popped in, I was using one.

Jane’s experience was less positive and reflected the way in which system components intentionally structured to provide feedback could be misused. When this happens, teachers do not benefit in the way it was intended. Jane offered her personal thoughts:

> Timely feedback would improve my practice for sure. Not to be negative, but I had my evaluation three weeks ago. It could be just because my principal is very overwhelmed but what her tactic this semester has been was to do all of the evaluations first and then go back and do the post conferences in the order that she did the evaluations. Now that she’s completed all the evaluations that she had to do for this semester, she’s got to do the post conferences.

New-to-MSD teachers in this study brought 4 to 20 years of experience to the position in this district and discussed the experiences of receiving feedback in the context of the level of expertise perceived to be held by the person delivering the feedback. Rose, a new-to-MSD teacher, offered a reflective look at how she has experienced receiving feedback:

> The quality of feedback from a veteran teacher helps me improve my practice. My mentor and principal give this amazing feedback. I never had a suggestion from them that I didn't take because they really know what is going on and they give non-biased feedback. What they've told me to do works. They are respected by kids and teachers, so I value and trust what they say. The verbal feedback and the discussions that come from evaluation are the most useful in changing my practices. I did have an experience where I got horrible feedback. It absolutely did not help me. It was all so critical. I was doing everything wrong. But the mentor I have had for the past two years is phenomenal. She watches me and sees what I am doing and she knows exactly how to make it better. Having her as a mentor has made me a better teacher so I know she gives quality feedback. She'll say things like, I notice when this happens, this happens. She notices the things I don't notice.

Experienced teachers identified using feedback to inform changes in instruction as an element within the existing system of evaluation that helped them improve their teaching. It was important to keep in mind that teachers classified as experienced

teachers in this study had been in MSD for four or more years. This had given them multiple years and multiple opportunities to experience the feedback structures embedded in the existing system and to develop personal attitudes regarding how these structures might be used to improve their individual practice. Actionable feedback, a term used solely by the experienced teacher group, was used to describe feedback perceived to be purposeful and effective. Katrina identified actionable feedback as something she could use right away and shared, "I like it when I can take it and do something with it in the next class. A good form of feedback is something that people can try." Sheryl, another experienced teacher, offered comments in the same vein:

> One of the things that improve teaching is being given feedback with something actionable to do and then you go and watch someone doing it and it helps even more. When that happens you get to see it unfolding in front of your eyes. It's not just theoretical, you got the theory and then you see the practice.

Other experienced teachers discussed their feelings when they perceived they were given "good" feedback. They affirmed each other's ideas of feeling personally and professionally challenged. Helen, a 20-year veteran, offered:

> I recently had an administrator who identified assessment as my refinement point and I never had that before. The idea of self-reflection, a clear target, a personal goal, all of the elements of formative assessment for me to apply to students was very powerful for me and I don't know that I would have gone in that direction if I hadn't had that particular feedback.

Sally, another long time veteran, expressed similar feelings regarding challenge and reflected, "I like feedback that makes me feel challenged. I like that I feel challenged at my school. I like that I am still learning." For some experienced teachers, though, it was not always easy to receive feedback. Some experienced teachers reported that "feedback I received once, has now become common practice," "after I met with my master teacher, I changed right away," and "the feedback I received made its way into everyday

practice." Others shared confusion or disillusionment with the way they received feedback. Kate shared: "I'm confused at the feedback I am getting right now. It's not making me think about how to change things in my classroom. The feedback is in the form of questions and I guess that is fine, but I can't answer the questions."

Kyle offered:

> I'm not the easiest person to give feedback to. I try really hard to do the things that are best for kids and it's so hard to get that right. So when I'm trying to get it right, and I feel like I've done a good job trying to get it right, and someone else sees that differently, my response is to be closed off to it and have a battle over it.

While only shared by one teacher in the entire participant group, we have to wonder whether the following sentiment was felt by a variety of teachers outside of the study. Anne reflected, "Sometimes it's hard for teachers not to take it personally and depending on how the observer gives the feedback it can feel kind of harsh sometimes even though I don't think that's what it's intended to be."

Dependency on method of delivery and quality of feedback. As alluded to in the previous section, teachers perceived feedback to be more valuable when it came from a trusted and well-respected source. Feedback from individuals not perceived to hold expertise in the area in which they are coaching was seen as less valuable. The way teachers reported feedback to be delivered included written feedback and face-to-face conversations. Teacher preferences emerged and were connected to teachers' feelings of how the feedback could help them improve.

Novice teacher, Ginger, was in her third year of teaching. She felt as if she "has just scratched the surface of education" and offered, "I think people are comfortable giving me feedback because I think I communicate clearly that I am comfortable receiving it. I'm totally willing to hear all feedback and suggestion from anyone." Many

teachers in this study viewed the feedback they received as a key piece to improving instructional practice but the way they received the feedback was perceived as integral to whether they were open to accepting it.

Within the new-to-MSD category, teachers shared a variety of experiences they have had with receiving feedback. Kristine offered:

> My mentor teacher comes in for 15-20 minutes and does a quick little write up with a few bullet points. Other times I've had other mentors come in and write paragraphs on what I can do to improve my teaching. I like the paragraph style better because it says you did this and kids responded this way. I am more of a visual reflection type of person so that is what I like. I like when I get to hear what I can do better. I also appreciate that I get some positive feedback first. If there is too much critical feedback I'm like, "Oh my gosh, I'm a terrible teacher," and I feel sad.

Valerie, an experienced teacher, has not had a positive experience as compared to Kristine. She offered:

> Getting feedback is usually a question approach, which is not my favorite because then I have to figure it out. Personally I like to get to the meat. What is it that I am going to work on? What am I going to do next? The feedback is less than I have had in the past. I know it's the mentor's caseload. You just get a little less feedback. I do get time to reflect upon the things I am working on but it's different than when I first started. Not as beneficial as it was.

It was important to understand the perspective of a master teacher as well since this group of teachers, usually within the experienced group, viewed the feedback component of the existing evaluation system through a different lens. They were the teachers who provided the feedback. Farrah shared:

> As a master teacher, I always leave written feedback with an invitation to contact me for more information or for a face-to-face conversation. People don't take me up on that very often. I wonder how much I should force that but then I think don't like being forced to meet with someone. So that's something I am working on myself, how to get people more involved in wanting that feedback. I try to match it to how they like feedback. It could be a quick conversation at lunchtime, just written, or sometimes face-to- face conversations.

Helen, also a master teacher, offered thoughts on making feedback relevant to teachers:

> Having feedback alignment between the master, mentor, and principal really works. People want the feedback aligned and connected to their work and to the messages they receive from all observers. What is not helpful is having un-aligned snapshots from multiple people who aren't necessarily communicating with each other or the people they are observing.

Receiving actionable feedback from a trusted source was a component of the existing evaluation system teachers perceived to influence change in their instructional practices. When the feedback came from a person not perceived to hold expertise in the content area being observed, teachers were less likely to implement the suggested change.

Differing perceptions of accountability for using feedback. Teachers did not perceive a strong link between feedback provided and expectations that feedback would be put into practice. Teachers lamented that a feedback loop in which they were observed applying feedback to receive new feedback was not prevalent. Brooke, a novice teacher, noted that she received feedback she could use in the future but did not feel accountable for using it. Rose, new-to-MSD, reflected, "There is no expectation that I will use the feedback . I think if I didn't do any of the stuff, none of them would have a problem with it but I value all of their feedback and so I do what they suggest." Within the experienced teacher group, the same reflections prevailed. Sage shared, "They never come in to see what I have done with the feedback." Sheryl expressed a similar sentiment: "I don't know that anyone comes back and says, 'How did that feedback go? What did you like about it? What didn't you like? What do you need to do to further improve?' I think that piece is missing from the current structure."

From the perspective of an experienced master teacher, feedback was not offered often enough. Xavier shared his experience:

> I only get feedback during a mentor, master, or principal evaluation. The observations have a standard pre and post conference format in which masters, mentors, or my principal informs me of reinforcements and refinement. Outside of that I don't get much feedback. As a master teacher I don't think people get a ton of feedback. I don't make it into classrooms as often as I'd like.

Joanne sums up the prevailing idea: "We need the feedback and then we need the mentors and masters to stay close and follow through." Joanne's comment illustrated teachers' interests in ongoing, continuous feedback.

Effect of critical comments following positive comments. Specifically in the novice teacher group, teachers expressed a need for critical feedback to be delivered on the heels of positive comments. These teachers agreed it was much easier to hear something critical when they also knew they had done something "right." Valerie and Mary clarified this position and shared their personal experiences:

> So my principal started with this really positive thing and then said, "The thing I want you to work on is…" and I was ready because he had just said this really positive thing and I felt so good about it because I knew what I had done well. Having up front positive feedback made me more ready to accept critical feedback or feedback about things I needed to change. (Valerie)

> Every time I have had an evaluation there has been a focus on what I have done well. There are accolades on that and then I'm given something to work on for next time. (Mary)

The experienced teachers in the study reflected: "Teachers would be uncomfortable with feedback if it were presented in a negative way." Michaela, from her vantage point of experienced teacher and past mentor teacher, noted:

> Teachers need to be able to get feedback that will help them to improve instruction, you know, positive suggestions. I think the reinforcement/refinement idea is a good one as far as telling us something we're good at, at the same time as telling us something we need to work on.

All of the comments regarding feedback were derived from the question, What do you think improves instructional practices? While other mechanisms were identified to be at play in terms of improving practice such as professional development, self-reflection, and watching other great teachers "do their thing," feedback stood out as the most common and most influential aspect of the existing evaluation system in terms of changing instructional practices. The existing system includes multiple opportunities for instructional leaders to provide and for teachers to receive feedback. The most common system structure for feedback mentioned by teachers was the post-conference. Without direct follow up or accountability for using feedback, teachers were left to their own devices to try new ideas. They were left to reflect independently about success or failure as they applied the feedback they were given. While most teachers expressed interest in receiving feedback, the willingness to apply the feedback was perceived to depend on the quality of the feedback, the method of delivery, the level of expertise held by the person providing the feedback, and the use of feedback loops in which a teacher received follow up observation in order to receive additional feedback on their interpretations and implementation of suggested practices. Even with all of these things in place, the quality of the relationship a teacher has with his or her mentor or master teacher can influence whether feedback is put into practice.

Quality of Relationships

The existing evaluation system includes a coaching component in which teachers are assigned a master or mentor teacher to provide them with feedback through a collegial relationship that develops over time. The perceived quality of the master or mentor teacher along with the perceived quality of the relationship between the teacher

and his or her master or mentor teacher were identified as elements that influenced changing practices. For some teachers, there was the perception that there would be no influence on changing instructional practices if the relationship was not positive. The perception that having a coach was ineffective was also mentioned. Jane, a current mentor teacher, summarized the different perspectives she has encountered:

> Some teachers like it (mentoring) and some do not. And I don't know if it's just my experience or if it's an overall thing but I feel like the newer teachers are really willing to soak up, like who I'm co-teaching with is a newer teacher to our district and she's willing to try different things and I'm hearing that through her evaluations. I'm noticing things that we've been working on and so that's really good. And there are other teachers who have been there for a while who, when I walk in their room, they're like, "What are you doing here?" "I'm like, I'm your mentor teacher, I'm your coach, and I've sent you five emails about picking a day and you didn't pick a day and so I'm here." I feel like everyone should have mentoring whether it's bi-weekly, weekly or…however often. But some people are choosing not to do it and I don't think there is any consequence for that.

Rose and Kristine, both new-to-MSD, valued the master/mentor teacher structure and remarked that the regularity of coaching visits helped them improve their practice. Both teachers found value in the positive comments and suggestions for improvement offered by their mentor teacher on a weekly basis. Rose commented:

> I value the master mentor program. I think if we ever lost that, we'd lose a lot. I value having people coming into my classroom regularly. My mentor teachers aren't judgmental. They talk to you as if what you are doing is not wrong but they give you suggestions on how to make it better. They don't belittle. They don't make you feel like you are a bad teacher. They have positive suggestions to make. I wouldn't be afraid to go to them with anything. They are very approachable.

Kristine held a similar perspective and offered:

> Having a mentor teacher come in every week has really helped my practice because it's overtime, it's not just once. My mentor comes in every week. She always asks me what I am going to teach and then she gives me a little feedback. You did great on … and then something to improve on.

Miranda, an experienced teacher, explained the way her relationship with her master and mentor teacher helped hold her accountable for the coaching suggestions they made:

> The biggest thing is just having someone come in to check up on you. Knowing someone is coming in to check on you makes you feel more motivated to implement the parts of the rubric that we are supposed to be. It makes it more accountable to have a check in. Having masters and mentors come in and check on you does help because if I know they are coming in I try to do it but if I know they are not coming in I might not take the time to do it. Having them come in does make a difference.

Brooke, another experienced teacher, commented on the changes she has seen in herself as a result of the master mentor relationship and remarked:

> I am such a better teacher than I was when I came here. Part of it is the master mentor system. Everyone is completely open to having people come in and see what they are doing. At some point we may stop caring if no one else cares about what we are doing in our classrooms.

In terms of relationships, Lucy shared the positive experience she has had:

> Having a strong mentor teacher that has a connection to you and that works with you in ways that don't turn you off or upset you but are absolutely honoring you as to who you are and what you are capable of. Honoring your practice, really. If you don't have the relationship, people are going to sit there and say, "OK, OK" but then they are going to do whatever they want. If it's not meaningful, if it's not purposeful, and if it's not collaborative, then it's not worth doing.

Taylor, a novice teacher, viewed the relationship she has with a master or mentor teacher as "what is expected." She commented:

> I'm going to be a teacher and this system is what's expected. I don't have a choice so I'm not going to complain about it. I think my classroom should be an open book and you can come in whenever you want to look at it. I want them (masters and mentors) to watch what I am doing. It makes me better.

Perceived value of master and mentor teachers. Some teachers did not see the value in having an assigned master or mentor teacher to serve as a coach for their instructional practice. All participant groups included individuals who communicated a

level of distrust for the mentor/master system of coaching. Sage, an experienced teacher, was very clear with her statement: "Having a coach is not effective for me." William put into words a common perception: "The mentors get all the knowledge from visiting other classrooms but they do not disseminate it." Other teachers expressed concern for different reasons. Valerie, a new-to-district teacher, explained that she is "not sold on the mentor yet" and continued:

> I see the potential value but it hasn't worked for me yet. I haven't had much success with the coaching conversations. The people who come into my room and get to say what level I am at, I think I should be able to go into their room to see where they are at. I would love to know what kinds of scores the masters and mentors are getting and why they are appropriate to come and watch and score me. Before they decide how I am doing, I want to know how they are doing.

Frida, although an experienced teacher, reflected on how nervous she felt when her master or mentor teachers were in her room. She explained that she "rushes through things and tries to fit everything in even if it does not work with the lesson." She perceives her mentor as not having the expertise needed to coach her. Frida disclosed that when approaching the mentor with questions, the response received was "You'll have to ask someone else." This was in contrast to the perception held of her master teacher whom Frida said "gives lots of suggestions for things I could actually try."

Perceived competency of the master or mentor teacher. Many teachers took the time to describe their idea of a competent mentor or master teacher. Sheryl explained her perception that if masters and mentors are doing their job well, "they should be in your room two or three times for large chunks of time, more than 20 minutes, and be giving you feedback." She elaborated, "I like having a scheduled dialogue with a trained observer. The key is having a very transparent and explicit structure. A coaching structure that does not ever say that you are a bad teacher, but rather says, I can improve."

Casey expressed her perception of a competent master or mentor teacher as someone who is

> constantly learning new things and seeking to improve themselves, not thinking they know everything….. It is someone who is open to making their own teaching practices better and who is open when you talk to them, listens to your perspective on why you did something or why did not do something and your reasoning behind it, and does not just say, "Well, it has to be done this way."

Katrina, an experienced teacher and also a master teacher, perceived that the relationships she has with the teachers she coaches is partly dependent on her being a classroom teacher also: "They know I am a teacher too, so that makes it huge. I cannot imagine being a master teacher, giving feedback, and not having my own 30 kids in the classroom. I do not know how I would be believable."

Teachers in this study made a point of talking about the interactions they had with their master or mentor teacher although they were never directly asked about this relationship. It was clear based on participant responses that the relationship between teachers and their master or mentor teacher had an impact on whether teachers valued the instructional feedback they received and also on how they felt overall about the existing evaluation system. Participants expressing a strong, positive relationship with their master or mentor teacher shared positive feelings about growing and improving their instructional practices through the elements of the existing system. Participants who shared feelings of "ineffective" relationships conveyed a general lack of trust for system elements through comments such as " having master and mentor teachers makes class sizes too big," "the masters and mentors have a wealth of knowledge from visiting other classrooms but it is not shared collectively with staff," and "all I heard from my mentor

was that my objective was not written on the board and I thought 'why does that matter' and then I stopped listening."

Effect of the Professional Practices Rubric

Often when the evaluation system was mentioned in MSD, teachers automatically thought of the Professional Practices Rubric (PPR). Some referred to the PPR as a handbook or manual, while others referred to it as a guide. Regardless of the moniker, the PPR was viewed as an instrument for promoting teacher instructional improvement. Teachers perceived the evaluation system to rest on the highly descriptive criteria found in the PPR. For teachers at any stage of their career, the majority of coaching took place around these instructional indicators. From these indicators, teachers refined practices in an effort to improve student achievement. Teachers agreed, whether novice, new-to-MSD, or experienced, that the PPR served as a guide for best instructional practices and that it had changed and improved their teaching practices over time. Although commonly perceived as an essential element for changing instructional practices, the PPR was not viewed as perfect. Teachers perceived the PPR to be "overwhelming" and "impossible to meet" as well as a poor indicator of professional practices for particular groups of teachers, specifically special education teachers. An additional criticism was the use and implementation of the PPR differed between the instructional leadership teams at different buildings.

Novice teacher Nancy reflected on how she thought the PPR was "just a performance review"; as a professional, she "thought she should read it." But then, "Holy smoke!" Nancy remarked, "Now I look at it as a guide and think about how it is going to help me." This perception was common regarding the value of the PPR. In all

teacher participant groups, teachers remarked that the PPR had "changed their practice," "improved their practice," and "made them a better teacher." Also included in these responses was overwhelming agreement that the indicators and criteria in the PPR were things you "have to do." New-to-MSD teacher Valerie shared:

> I feel like the things that are in the PPR are things I learned that you have to do. The PPR is a good indicator of good teaching. Any type of way that you can look at a document that's sort of like a report card and see where you're at, I think that's useful. There are things in there that you have to do. They are good reminders especially once you get them and you go over them. It's important to try to follow it. It's hard to remember everything, for every lesson, but it's impactful for sure.

Kristina and William, also new-to-MSD, echoed the sentiment with "the specifics about what good teaching looks like are in there" and "the rubric tells you all the things you should do as a teacher." Susan, another new-to-MSD teacher, noted, "My results have changed because I have tried to get to know the rubric. A friend of mine from another school said that the ones who score really well are the masters and mentors because they know the rubric inside and out." This comment directly aligned with a remark made by a current master teacher:

> The PPR is a pretty effective indicator of my teaching. I now know the PPR pretty well and I think about the things that are in in it and I know it has helped me become a better teacher and helped the achievement of my students. The rubric has made me a better teacher and as a master teacher allowed me to make other teachers better teachers.

Within the experienced teacher group, comments remained similar. Joanne offered, "The PPR made me a better teacher. All teachers are better because of it. If we didn't have the indicators for presenting a lesson or lesson structure, teachers would not have something to base their teaching on." Katrina offered, "The rubric, well those are indicators of effective teaching so if you can look at the rubric, and look at the range of

scores, you can use that to guide your instruction." Even with an overall perception that the PPR could help teachers improve their practice, teachers wanted clarity on the indicators "before they can be expected to do it." Both Brooke, an experienced teacher, and Susan, a new-to-MSD teacher, shared: "If the rubric is about good teaching, teachers need to be very clear on why it is good teaching before they can be expected to meet those expectations."

Although the PPR was seen as an element within the existing system of evaluation that helped teachers improve instructional practices, common criticisms were mentioned by teachers in all three participant groups. The sheer length of the PPR was identified as a deterrent to its usefulness. Novice teacher Ginger said, "The rubric is overwhelming and it covers so much, every lesson, every day. It is unreasonable and unrealistic because nobody has that much time even if you get into a routine where you know how to plug everything in." New-to-MSD teacher William offered, "It is a lot for a person to keep track of." Experienced teachers held strong beliefs regarding the length of the PPR and offered similar comments such as "the rubric is impossible to meet." Frida, one of the experienced teachers in the study, commented: "The whole rubric is a little much. It's kind of overwhelming." Helen, also an experienced teacher, expressed feelings of futility in using the rubric to guide practice: "You have to do 147 things in a 55-minute period and that's just silly because you could not observe all those things in 55 minutes any more than you could demonstrate them." Kyle also commented on the length of the PPR and offered: "Our district is saying here are 133 steps you can do in 50-minutes to be effective." Although counting the indicators and criteria in the PPR yielded different numbers, participants focused on numbers as evidence that the PPR was

too big. Sally also shared her point-of-view and although she recognized it is "good to have those things written out so you can do them and be a good teacher," she also felt that "there is too much in there" and wondered how anyone was supposed to "do all of those things in every single lesson." Holly shared her perspective of the growth she had experienced as a teacher who was taken aback by the length and specificity of the PPR:

> When I arrived here I felt like I wasn't recognized for the skill that I brought to the table. And I think it was just that my skill didn't match all of the little indicators in each domain on the rubric. There was a lot of that going on. Now, I keep a lot of those domains, they're always in my mind. They're in my mind when I'm planning. They're in my mind when I'm with the kids and when I'm teaching. Predominantly, they're in my mind because I believe they're good practices. But I will say that sometimes I'm not…you know, there used to be a lot of spontaneity….not wasted time or things like that but like, "Hey, we're going to make this project today" and just get out the construction paper and maybe the kids were just having one of those days and maybe we decided to do something fun before we went to break and I don't really feel like I have latitude to do those sorts of things.

In addition to the length of the PPR, teachers were critical of the fact that they did not perceive it to measure what they did as teachers. Susan, new-to-MSD, clarified the perspective of the special education group:

> I don't think it measures very well. It measures my instruction well, but instruction is only about 1/3 of my position. Part of my day is direct instruction, part is support instruction and a lot of it is paperwork and making sure kids' needs are being met in their classrooms. I don't get feedback on 2/3 of my job.

Frida and Miranda agreed with this statement and shared:

> I don't think it measures my teaching very much because what I do to make it work is show a lot of extra things. It's not really what I am doing. It's a tough rubric as a special education teacher. (Frida)

> It measures my teaching pretty well but in Special Ed, so much of the job as a teacher is not just teaching. We are not observed during IEP meetings or keeping track of all the paperwork. On the PPR the professional part is 5% and that's really more than 5% of our job for sure. I feel like I hear from a lot of Special Ed teachers that the rubric doesn't really incorporate everything they do as a teacher.

> It looks at the instructional part in front of kids but it doesn't really look at everything else. (Miranda)

Another common criticism of the PPR was although the PPR was identified by participants as an accurate measure of teaching practices, it was perceived to be limited because it did not measure relationships between teachers and students. Holly explained:

> I just feel you can tell so much with rapport and even looking at student achievement and how kids are excited to go the principal's office to show their journal response in the morning. Those are the things that I feel like make it good for kids and I think if kids are showing those qualities then I think that there are probably some good things going on. That's not exactly touching on the evaluation system.

Kyle added: "The research-based parts of the rubric measure my practice, but it doesn't measure how well I interact with kids, reach kids, connect with kids. Our rubric doesn't even begin to address these things and they are the things that really matter."

A final criticism centered around the use and implementation of the PPR. Although the PPR was viewed as a standard measure of teaching practice throughout MSD, it was not understood or implemented in the same manner across all schools. Experienced teachers, most of them members of their school's instructional leadership team, weighed in on the differences. Katrina elaborated:

> Peoples' knowledge of the PPR can be very black and white and that can be limiting. Instructionally the PPR does a good job. You can get a feel for your strengths and weaknesses and your areas that you are comfortable in, but there are some spots that are a little difficult, some of the indicators are a little slim. For example, there is one indicator where there is literally no difference between scoring a three and scoring a five. There has to be a lot of continuing education for people in the system. Even if it is a fair tool, when the people doing the evaluation have more knowledge than the people receiving the evaluation it is not a fair tool any more.

Sally shared her memory of when the evaluation system was initially introduced:

> When we first started this system, new teachers were educated on the indicators and what they meant and for some reason new teachers are not taught the rubric

> anymore. I think that some teachers feel frustrated because the first few observations get really low scores and they didn't see that coming and they feel defeated.

Helen, also an experienced teacher and member of her school's instructional leadership team, shared the components she thought were necessary for teachers to improve within this system: "If everyone is working with a clear rubric and clear criteria and you have a decent relationship then you can begin to coach on the individual indicators and maybe begin to make some progress."

The PPR was viewed by participants in this study as an effective tool for guiding and improving instructional practices but it was also viewed as overwhelming. None of the participants in this study offered suggestion for changing the PPR to make it shorter or more condensed; however, participants did discuss adding criteria to it to capture areas they viewed as underrepresented. The majority of responses from participants identified the PPR as a tool that "guides," "informs," and "structures" teacher instructional practices. Special education teachers participating in the study found fault with the criteria on which they were measured and other teachers viewed the PPR as limiting since it did not measure the quality of the relationships they had with students.

Effect of Modeling

Telling someone what to do and showing them what to do were two very different experiences teachers reflected on as a means of improving their instructional practices. Teachers viewed modeling as integral to their growth. Some teachers perceived the existing system to provide avenues for facilitated peer visits and/or modeling by exemplary teachers. Teachers in this study believed seeing great teaching in action was critical to their development but it did not happen enough.

All teacher groups within this study identified modeling as a strategy used by master and mentor teachers to demonstrate the suggestions they made for teachers to improve their practices. Ann, a novice teacher, suggested that although master and mentor teachers were "always modeling and teaching different strategies," that was one thing she would have liked more. Ginger, another novice teacher, expanded on this idea:

> When it comes to professional development it needs to be more based on actual implementation and modeling and not that sort of, "sit and get" lecture-based learning. I think 21st century learning needs to be a lot of modeling, a lot of "let's see what this looks like in real life." Let's see somebody teach a lesson with a particular strategy to show how it's effective, and why it's effective, and how you do it. How you actually implement it. So, I guess for me, because I am a visual learner, I like to learn through experience and I don't know how to do that unless I get into other rooms and see it.

Valerie, new-to-MSD, offered a clear way for her to benefit from coaching: "I feel like there should be a lot of modeling. I think taking a piece of the rubric and saying, 'Okay, this is a piece of the rubric' and taking small steps and saying, 'I'm going to show you how to get a five on this.'"

Experienced teachers focused the most on opportunities to experience modeling of instructional practices as a reliable method for improving practice. Unfortunately, they lamented, "It does not happen as often as it should." Brooke expressed that her confidence in her own ability would increase "if someone came in and modeled what various levels of questioning might look like depending on the subject they want to see that in." She further elaborated: "We do not get as much direct modeling as I would probably like. I want to see teachers do things well so I can see how it's done." This was an outlook shared by many teachers in the experienced teacher group even though individually they had been teaching from 4 to more than 20 years. Xavier offered, "Modeling does not happen as much as it should. It is not just presenting it to teachers

but also modeling it in their classroom and showing them what it looks like and can look like in their room." Sage, an experienced teacher, did not believe she benefited from coaching: "Watching other people teach is a great way to improve your own practice." Farrah, an experienced teacher and also a master teacher, explained her outlook on modeling:

> We have to model it for teachers. I think if teachers know what those behaviors are that will help students, and if they have been explicitly shown how they can work in a classroom, then they can take that into their own practice and it does help student achievement.

The existing system of evaluation includes an expectation for master and mentor teachers to provide direct modeling of their feedback and suggestion or to facilitate classroom visits for teachers. When participants in this study were asked, "What improves classroom instructional practices?" the subject of modeling came up unprompted in more than 50% of participant interviews. Seeing great teaching in action was viewed by participants as "proof" or a "concrete" example that the instructional suggestions really worked for students.

Effect of Personal Integrity and Self-Reflection

Teachers spoke about possessing an inner motivation or a particular level of personal integrity enabling them to maximize the benefits of the existing evaluation system and access avenues for improvement. Teachers also discussed a perceived need to be open to coaching and feedback for it to "work" and they attributed this openness to personal characteristics brought to the table by the teacher. Integrity, reflectiveness, and internal desire to improve were qualities specifically identified by teachers in this study to be supportive in terms of helping them benefit from existing structures within the

evaluation system. More than half of the teachers interviewed in this study made reference to personal qualities they felt they possessed that enabled them to "make the most" of the evaluation system. "It's going to take me and my integrity to get students where they need to be," stated Nancy, a novice teacher. "It takes grit to be a reflective person who implements reflective ideas," agreed Kristine, a new-to-MSD teacher. These comments were not far from the perspectives held by experienced teachers who identified inner motivation, self-direction, receptiveness, desire to improve, and willingness to reflect on feedback and apply it as characteristics that helped them grow and improve in their profession.

Teachers in this study agreed that the structures existing in the system could be beneficial in terms of supporting change and improvement in instructional practices but only if the teacher was willing to accept feedback, reflect on teaching, and be willing to change things as needed. Sally, an experienced teacher, offered the following perspective:

> I am so receptive. I'm eager to change and learn and grow because that's my work ethic. If I need help, I'm going to go and get it. I'm not going to wait for this stuff to come to me. Feedback could potentially have a huge impact on changing practices but it depends on the receiver. Some people do not appreciate feedback and they push back. I think that's about the personality of the teacher. My own teaching has improved because of the specific feedback. But it's because I take the feedback genuinely and really respond to it. They give me the tag, the path, and then I take it the way that I need to.

Kyle offered his own take on the characteristics he perceived to be essential for improving instructional practices:

> Passion, relationships with kids, desire to improve. That's what it takes. Personal drive that's intrinsic. Wanting to be better for the sake of being better, wanting to be better at instruction for kids, wanting to be better so kids can improve and be passionate about learning too. It's mostly intrinsic.

Tess added another characteristic, her ego, stating, "I take my feedback and I apply it. Part of it is my ego I want to be well respected and if this is what my evaluators think is good then I am going to make sure I meet that bar."

The level of intrinsic motivation possessed by individuals was viewed as integral to teacher growth and development within the existing system of evaluation. It might be inferred from the participant responses that everything one needs to improve their practice exists within the system if one possess qualities such as openness to suggestion, willingness to make change, intrinsic motivation, and skills of self-reflection. Sheryl summarized her thoughts on the existing system of evaluation as related to the qualities teachers viewed as necessary in order to benefit from the system:

> Some people who have been doing it for a long time think the evaluation system is penalizing because they have their way of doing things and they don't want to change. They are not willing to grow or break apart the indicators for teaching. It would be hard to want to participate in system activities if you knew your methods were not looked upon favorably.

Connected to the belief that teachers hold the power to improve within their personal attitudes were responses to the following statement: all teachers want to be great but do not know how, or what, to do to get better. This question comprised the final discussion point in the interview process. Responses to this statement were more varied than any of the other interview questions or topics. Evidence that teachers were passionate about their profession resonated in every response from disbelief that a teacher would want to be anything but great to lack of understanding about how "anyone could not know what to do or where to look for help." Miranda shared, "Most teachers want to be great, but in many places teachers do not have an outlet to get better." Casey said, "For those teachers who are receptive and eager to learn, anything you can give them to

make them a better teacher, they will take it and run with it." Teachers were divided in their responses to the statement. Farrah shared:

> I could see that being true for some people. There are just different kinds of learners. There are some that are self-directed and are able to find what they need to make themselves better but I think a lot of people need the coaching and the feedback and people helping them and giving them advice on how to get better.

Tess offered a similar thought:

> The statement resonates with me because for teachers there is such a personal attachment to performance that I think it is not that they do not know what to do, it is that they do not know how to do it. If teachers are not hitting 4's and 5's they think they are not a great teacher and it is an ego hit because they think they are really doing a good job. Some teachers do get past it. The ones who sit down and really reflect on the good, the bad, and the ugly of their lesson are the ones that are constantly tweaking their instruction in their head and as they do it. The ones who think "that was great" with no reflection are usually the ones who are unhappy with the feedback they get.

But, Sheryl was very clear about her thoughts on the statement:

> I do not think that's true at all. Our whole district is deep into professional development so you must be asleep if you do not know what to do. Our district knows what we have to do to get results from kids. If you do not know what to do then you are not paying attention. I cannot imagine that you would not know what to do but if you do not know how to apply it, that would be a conversation to have with your Instructional Leadership Team. I mean, there are so many people to ask and if you do not, then shame on you.

Katrina addressed both sides of the statement:

> I do think teachers want to be great. I don't think there are many teachers who are like "whatever." There are some to whom it does not apply but trying new things is scary for some people. So, they might know what to do and how to do it, but they might be intimidated. I think they know where to go for resources. Some are a bit more hesitant and it might be a trust issue or something. Or maybe they want to do it themselves because they are a perfectionist. They want to invent the wheel themselves.

William reflected on his experience over time: "What comes to mind is newer teacher versus more experienced teachers. I am at the point that I have so many things in my bag

but it's just putting them all in. I could relate to that statement more in my earlier years than I do now."

The responses to the statement were widely varied and represented individual perceptions teachers had of what it means to grow and learn and improve within the teaching profession. Joanne, in her final statement of her interview, encompassed many of the themes mentioned within this study:

> I don't believe that statement. I don't think there is a teacher in this district who could say, "I don't know what to do to be a better teacher." We have resources. We have the rubric. We have master and mentor teachers. I hope that all teachers would know that. They should not have to go out and look for it on their own.

The statement, All teachers want to be great but do not know how or what to do to get better, struck both nerves and chords for the participants in the study. It prompted teachers to reflect on the elements existing within the MSD evaluation system and the value of the elements in terms of improving practice. An illuminating comment reflective of the current political state that inspired teacher evaluation reform was made by an experienced, yet new-to-MSD teacher: "Some teachers will never be great even with a ton of coaching and support."

Summary

Chapter IV reported the responses of 30 teachers regarding their thoughts on how the evaluation system existing in MSD might influence their classroom instructional practices. The scrutiny of more than 30 hours of one-to-one interviews illuminated a range of perceptions. It became clear through analysis that perceptions existed on a continuum from positive experiences within the existing system to negative or non-impacting experiences. Some teachers expressed, "I am a better teacher as a result of this system" and others countered with "the system does nothing to help me improve my

practice." Within each emerging theme, distinct elements of the existing system were identified. Each was perceived to be instrumental in terms of improving classroom instructional practices. Feedback, modeling, and a positive coaching relationship with a master or mentor teacher were consistently identified as elements within the existing system that exerted positive influence on teacher instructional practices. Self-reflection also emerged as an element teachers perceived to influence their practice and was closely related to participants recognizing their personal levels of intrinsic motivation.

An opposing viewpoint, derived from questions surrounding whether teachers perceived the existing system to restrict professional growth, yielded the observation that some teachers viewed the existing system as limiting in terms of improving classroom practices. Some teachers perceived that it did not measure the quality of relationships they had built with students or the level of humor or engagement they brought to the classroom. Teachers within the field of special education consistently shared that the professional practices rubric (the tool used to measure, guide, and inform teacher growth in instruction) did not measure the job they did.

Most participants considered the evaluation system and its components of observation, feedback, coaching, modeling, and the PPR as highly useful and influential. However, other comments demonstrated a perceived lack of value in the system. Overall, components of the evaluation system were viewed as intentionally supportive for teachers. In a system designed for implementation in a standard manner, teachers illustrated through their comments that very different practices might be occurring across the schools in the district. Existing evaluation system practices might not be meeting the needs of all teachers.

Chapter V further synthesizes the research and situates the findings within the greater context of an evolving system of evaluation within the school district and the state of Colorado.

CHAPTER V

DISCUSSION

The purpose of this study was to identify perceptions held by teachers currently observed and evaluated on a particular set of instructional criteria existing within an evaluation system in one Colorado school district. Colorado's mandate for reporting teacher effectiveness, based in part on teacher evaluation criteria, has increased the urgency for districts to examine existing evaluation systems (Colorado Department of Education, 2014). Teacher perceptions of how evaluation practices might improve instruction and the way in which a system of evaluation informs change in teaching practices might lead to increased understanding of how an existing system could be modified to achieve its intended outcomes. The following research questions were addressed through participant responses in Chapter IV:

Q1 How do novice teachers feel the district evaluation system impacts their classroom practice?

Q2 How do experienced teachers feel the district evaluation system impacts their classroom practice?

As addressed in Chapter IV, a third, unintended question was addressed through participant responses: How do new-to-MSD teachers feel the district evaluation system impacts their classroom practice?

The conceptual framework for this study reflected the idea that teacher evaluation is a means to improving the educator workforce by providing specific, observation-based

feedback, follow-through specifically designed to meet teachers' needs at their current level of practice, and opportunities to provide teachers with information to move their practice forward. Little existing research has documented teacher perceptions of how evaluation processes improved their practices. For this reason, this study was significant. This research illuminated teacher perceptions of the mechanisms existing in one evaluation system that have the potential to begin building a foundation of research in this area. An objective of this research was to capture the teacher voice, enabling system designers to respond to the ways teachers feel the system of observation and evaluation impacts their professional growth.

In sharing the experience of observation and evaluation within the boundaries of an existing evaluation system, participants presented their understanding of the current system, the challenges within the system, and teachers' own unique perspectives on whether the existing system of evaluation benefitted their growth as professionals. Five overarching themes emerged from the data collected. These themes revealed the complexity of the system in which multiple mechanisms must work in coordination for teachers to realize the benefits.

Understanding the perceptions held by teachers of the evaluation system in which they performed their duties was needed. In this era of state educational reform and high accountability, it might be that the voice of the teacher is unheard above the din of legislation. This research gave voice to the perspectives of individual teachers. This voice illuminated aspects of the evaluation system perceived to be working and that teachers believed should be maintained. The research also identified areas teachers perceived to be in need of modification and change.

In this chapter, the researcher describes these perspectives within the context of the emerging themes. The researcher will draw conclusions based on the data, make suggestions for district-wide consideration, and identify possible paths for further research. A section containing teacher recommendations for change is also included.

Teacher Perceptions

Of the 30 teachers involved in this study, 93% valued the existing system and believed it had made them, or was making them, better teachers. In more detailed numbers, 28 of 30 teachers ranked the existing evaluation system above a "5" in terms of their perception of whether it helped to improve their instructional practices. In MSD, teachers viewed the evaluation system as comprehensive and included far more than just an evaluation score. Teachers in this study expressed understanding of the ways in which observation and evaluation were supported by other system structures such as master and mentor teachers, ongoing coaching, professional learning communities, and the PPR. Teachers in this study believed the evaluation rubric provided them with a guide for best practices and a road map for "how to teach." The evaluation system in MSD strove to include mechanisms for "growing teachers" as much as it strove to meet state educator effectiveness criteria. For the evaluation system to avoid being viewed as punitive, measures were included that were intended to provide support for teachers. An evaluation system that measured performance but did not give teachers clear direction for how to improve their performance would not have value according to teachers in this study.

As mentioned previously, 28 of 30 teachers viewed the existing system instrumental in improving their practices. Of the two teachers in the study who ranked the system below a 5, one of them ranked it a 3 for reasons actually supportive of the system. This teacher talked about incorporating the things she heard during her first evaluation, made suggestions she heard from her master and mentor teacher a part of daily practice, and no longer felt the evaluation itself was impacting changes in practice. The teacher benefitted from the PPR early on and used it to inform instructional practice. The other participant ranked the system a 2 and disclosed that she "did not know much about the system." She offered that she did not have enough time to pay attention to the evaluation system or the criteria the PPR was measuring even though she thought the rubric was good. Since she was a new-to-MSD teacher, it illuminated a need to make connections for teachers, help them understand the power behind the professional practices that comprised the PPR, and immerse them in a system not only designed to measure instructional practice but to provide support in areas in which teachers need help.

Although the existing evaluation system was supported and overwhelmingly viewed by teachers in this study to help improve their practices, distinct areas within the system teachers were identified as essential. The following section discusses each of the themes identified in this study. As stated by one participant, it was not the evaluation that impacted practice but the professional development that resulted from the evaluation.

Effect of Feedback

Within the existing evaluation system, an expectation was teachers would receive feedback on their practice by providing them with accolades on their successes and

evidence that their practices seemed to be working. Teachers would also receive feedback on potential areas for modification or change. Although this was the intent, many teachers reported they did not receive feedback often enough. It was clear from this study the feedback teachers received most often was a result of their evaluation observation and was included in their post-conference. Teachers believed feedback limited to their evaluation was not enough to help them improve. Teachers desired feedback in an ongoing manner. They wanted to be told what they could do to improve and they wanted someone to follow up with them. When teachers received feedback without anticipating any follow up, they did not feel compelled to implement it. They wanted to hear how their interpretation and implementation of the feedback looked to an outsider visiting the classroom. They wanted to be successful in their practice and wanted someone to care enough about the feedback they provided to make sure the feedback worked. They wanted someone to provide new feedback based on continued observation.

The existing system included mechanisms for doing just this but when system structures were not implemented in a uniform way or in a way that felt good for teachers, the outcome was a lack of belief in the system. For example, when a principal developed a schedule to complete all observations within a given period of time and did not plan post-conferences or provide feedback sessions until all observations had been completed, that was unfair to teachers and disrespectful of the system. Teachers did not perceive feedback shared three weeks after an observed lesson to have the same impact on changing instruction as immediate feedback. The professionalism of the teacher was devalued when he or she was not provided with timely, actionable feedback.

Teachers in this study made it clear they valued feedback when it came from an experienced source, preferably with expertise in the content area of the teachers they were coaching. This related to the selection of individuals for master and mentor positions. Since some of the participants in this study were currently, or had been in the past, master or mentor teachers, the researcher gleaned multiple perspectives from individuals in these roles. Some mentors expressed they felt they were not taken seriously by the people they coached. They noticed their comments and suggestions were met with nods but they did not feel their comments were valued. Teachers in the study explained they did not value feedback when it came from an inexperienced source or a person who did not know their content area. Special educators in the study shared this concern consistently. Collectively, they did not feel they received instructional feedback from people who understood special education often enough. When hiring or assigning teachers to master and mentor roles in the future, it would be important to select individuals with intentionality in terms of meeting the needs of the staff they support. Certainly budgeting restraints, locational challenges, and a host of other issues would come into play but teachers in this study reported having an experienced teacher with expertise in their content area was critical to whether or not they would accept and implement the feedback they were given.

Quality of Relationships

Quality of relationships emerged as an element teachers perceived to have a strong impact on how they viewed the evaluation system in which their teaching was measured. This might be another area connected to the actual hiring or assignment of teachers to master and mentor roles. In MSD, the qualifications for master or mentor

teacher did not include interpersonal relationship skills. From the evidence gathered in this study, the ability to relate to other individuals was a critical aspect of whether or not a master or mentor teacher would be successful. Teachers expressed the need for a positive relationship with their master or mentor teacher. When one existed, they actually wanted to improve, change, and implement the feedback suggested. Teachers felt less threatened when they had a positive relationship with their coach; they found the feedback to be more aligned with their practice. The concept of the existing evaluation system was to grow teacher practices through coaching based on evaluation and observation, yet coaching is a difficult skill to acquire. Master and mentor teachers who do not innately possess skills for coaching teachers to their personal best cannot be expected to acquire these skills through trial and error. They need to be specifically taught. The value of coaching relationships would increase exponentially if teachers thought the feedback and coaching they received were of high quality. An individual who is well respected among his or her peers would be more impactful over time than a person who has expertise only.

Throughout the study, teachers mentioned not being sure their master and mentor teachers were actually that skilled. They wanted to know the scores achieved by their master and mentor teachers to validate the coaching and evaluation feedback they provided. The teachers thought they would be more likely to accept feedback provided by a master or mentor teacher if those individuals demonstrated high quality, impactful instruction. As one master teacher put it, “I don’t see how I would be believable if I didn’t have a class of 30 students of my own.”

Distrust in the master and mentor system could eventually break down the system of coaching and feedback. Care must be taken to place the appropriate individuals in the appropriate roles. Somehow, the inclusion of a prospective master or mentor teacher's status within the school in terms of interpersonal relationships and quality of instruction must be considered when hiring people for these roles.

Effect of the Professional Practices Rubric

The PPR was identified in this study as the system piece teachers most often thought about in relation to the evaluation system. Although perceived to have a strong and positive effect on instruction, it was also perceived to be too big and overwhelming and quite impossible to meet. It was also not perceived to measure practices and specific job responsibilities within certain instructional positions. The issues with the PPR were not that the PPR asked teachers to demonstrate inappropriate or ineffective practices, rather there are too many practices to keep track of. A common perception and complaint was all items could not be observed in one lesson. This was addressed in the past by MSD by terming some practices as snapshot indicators, items that should be seen in every lesson; whereas other practices might be more discreet or specific to a particular lesson or learning target. For example, the type of thinking students used to solve a particular problem or the cooperative grouping structure students participated in to create a product or solve a problem might change depending upon the lesson. It could be inferred from teacher comments throughout the study that the PPR needed role-specific adjustment. All of the special educators in the study found the PPR to reflect poorly on their actual practices. Thought needs to be given as to whether the criteria for measuring performance adequately measures the job teachers do. While the indicators for

instruction were not disputed by special educators, more emphasis was based on instruction than on the responsibilities for legal paperwork, parent communication, and school-wide communication of interventions. Restructuring the criteria within the professional responsibilities area of the rubric or differentiating the weights allotted to each domain of the PPR might be ways to address this concern.

Effect of Modeling

Modeling was embedded in the coaching component within the existing system of teacher evaluation but was not perceived to take place often enough to influence change in teachers' own practices. Teachers wanted to see how instruction unfolded for others they perceived to have expertise. No matter the participant, school, role, or responsibility, modeling was not perceived by teachers to be currently used as an effective coaching strategy. The avenues for providing coaching, feedback, and modeling existed but were perceived to be underutilized. Teachers desired to see best practices in action more than they wanted to talk about how they might look in a classroom. The data collected and analyzed throughout this study implied this was not an issue unique to one or two schools. When teachers see their master and mentor teachers modeling the feedback they are suggesting, there is a true opportunity to develop trusting relationships with the goal of bringing about professional growth.

Effect of Personal Integrity and Self-Reflection

From this study, the researcher gleaned that a teacher's own feelings of efficacy were influenced in both positive and negative ways. The theme of personal integrity and self-reflection as mechanisms for benefitting from the existing system and growing one's own professional practices were shared by a majority of the participants in this study.

The belief shared by participants was a teacher's willingness to reflect on his or her own practice and make instructional change based on feedback received could truly improve professional practice. When teachers in the study were asked to respond to the statement, "All teachers want to be great they just do not know how or what to do to get better," comments related to self-help and self-reflection were common. Teachers in this study shared the desire to be great and perceived themselves as reflective individuals who were motivated to improve. They believed they knew how to work within the system and used the elements it included to their benefit.

Conclusion

Some research within the last decade reported that studies of teacher evaluation practices showed little effect on improving teacher practices (Donaldson, 2009; Ellet & Teddlie, 2003; Peterson, 2000) but other research has begun to identify the benefits of well-designed evaluation systems (Donaldson, 2009; Kane et al., 2011; Kimball et al., 2004; Milanowski, 2004; Milanowski et al., 2005; Taylor & Tyler, 2012; The New Teacher Project, 2013). This research supported the idea that teaching could be improved through a well-designed evaluation system. Teachers in this study believed they were better at teaching as a result of the existing evaluation system. However, careful attention must be paid to teacher beliefs that it was not the evaluation alone that helped them to improve their practices. Rather, the combination of multiple system characteristics working in harmony helped them improve.

The model for teacher effectiveness and evaluation embedded within Colorado's Educator Effectiveness Act (2010) included many of the elements teachers in this study perceived as essential for positively impacting professional growth: clear and specific

teaching standards, trained evaluators, and a valid scoring rubric teachers could use to guide and track their skills and development. This research confirmed the importance of these elements for teachers. Through examination of the shared perspectives of teachers in the study, these elements surfaced as essential characteristics of an evaluation system teachers perceived could help them grow and improve their practice.

The purpose of this study was to uncover teacher perceptions of how the existing system of evaluation might impact their growth as professionals. Through the research, it became clear teachers believed this system helped them improve their classroom practices. The combination of a strong rubric that acted as a guide for best instructional practices and trained evaluators who are knowledgeable about their content and can develop strong interpersonal relationships were identified by teachers as the core of the evaluation system. It was not the scoring of instruction teachers perceived to impact their willingness to change and/or improve practices but rather the harmonious interaction among rubric scores, coaching feedback, and a teacher's personal integrity and willingness to implement the feedback they received.

The existing system, although strong in areas such as the professional practice rubric, was not perceived as constructed or orchestrated perfectly. While there are many benefits to the existing system, care must be taken not to lose sight of the importance of feedback and coaching in terms of improving practice. The assignment of scores to an instructional moment in time was not the piece teachers in this study believed would improve practice. "Chasing the numbers" and embarking on a quest for the elusive score of "5" was not perceived as essential to teachers in the study in terms of how their practice improved within the existing system. It was the cycle of observation and

feedback that led to improved practice. This cycle did not produce intended results unless the providers of the feedback and observation were perceived to be strong educators themselves. The benefit to teachers increased when the master or mentor teacher also possessed content knowledge in the observed teacher's area of instruction. According to teachers in the study, professional growth also depended on the ability of the teacher to self-reflect and implement suggestions. The components of the existing evaluation system would not have the intended effect if they functioned separately from one another.

While the term "evaluation system" connotes a system of measurement, the existing system in MSD reflected a model of professional growth more than a system of measurement. For a system of teacher evaluation to be perceived as beneficial for teachers, it has to provide a clear indication of instructional strengths with feedback for continued practice, identification of areas of relative weaknesses within a teacher's instruction, and specific suggestions for change. For an evaluation system to work for teachers, everything within the system has to work together.

Teachers made a variety of suggestions for change in the existing evaluation system. For any system that directly impacts teachers, their voices must be clearly heard and addressed. It was the hope of the researcher that this study would continue to illuminate areas within the existing system that were working and should be maintained while addressing areas that would benefit from change. As the state and the nation continue to try to develop systems that accurately measure teacher performance and effectively improve teaching practices for the benefit of students, it was important to hear these voices. These suggestions came from the hearts and minds of frontline educators.

If the system can measure performance, thereby influencing whether teachers have a positive impact on student achievement, the following teacher-made suggestions are worth considering for future modification to the existing system:

- If there is more focus within the system on feedback, coaching, improvement, and growth rather than on straight up evaluation, teachers would feel better about it.
- Choosing not to have any mentoring should have some consequences. If everything were implemented the way it is intended (feedback, coaching, pre conferences, post conference), each element would be equally effective in evaluation.

This research supported the decision made by MSD to maintain their existing evaluation system even as the state of Colorado takes steps to refine a statewide system for measuring teacher effectiveness and improving instructional practices. Mountain School District utilized a complex structure of professional development to help teachers enhance and improve their classroom practices. The acceptance of and belief in these structures expressed by participants in this study indicated the existing system was perceived to have value in terms of improving practices. However, this research uncovered potential areas for further study within MSD. For example, teachers expressed they were "better teachers" as a result of this system but did student achievement results reflect this perception? Teachers also expressed belief that the criteria in the PPR were examples of best instructional practices but were those practices in line with the level of rigor and current expectations for the teaching and learning embedded within the Common Core? Although the existing system of evaluation was perceived to be

beneficial, these questions provoked consideration for additional qualitative study regarding perceptions of how the PPR reflected current expectations for teaching. The research also pointed to a need for quantitative study in which teacher practices, evaluation scores, and effectiveness ratings are aligned with student achievement. Without this additional information, we might continue to put effort into a system that is well received but is not producing the desired outcome of increased student achievement.

REFERENCES

Bambrick-Santoyo, P. (2012). Beyond the scoreboard. *Educational Leadership, 70(*5), 26-30.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Battelle for Kids. (2012). *Global education study: Six drivers of student success--A look inside five of the world's highest performing school systems*. Retrieved from http://www.battelleforkids.org/initiatives/initiatives/global-education-study

Buddin, R., & Zamarro, G. (2009). Teacher qualifications and student achievement in urban elementary schools. *Journal of Urban Economics, 66*(2), 103-115.

Chingos, M. M., & Peterson, P. E. (2010). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. *Economics of Education Review 30*(3), 449-465.

Colorado Department Education. (2011). *Users guide for evaluating Colorado's teachers.* Retrieved from http://www.cde.state.co.us/EducatorEffectiveness/SMES-Teacher.asp

Colorado Department Education. (2013). *Implementing the model system for evaluating Colorado's educators*. Retrieved from http://www.cde.state.co.us/educator effectiveness/trainingtools

Colorado Department of Education. (2014). *Educator effectiveness: A Colorado priority*. Retrieved from http://www.cde.state.co.us/EducatorEffectiveness/index.asp

Colorado Legacy Foundation. (2013). *Colorado integration project*. Retrieved from http://colegacy.org

Cosh, J. (1999). Peer observation: A reflective model. *ELT Journal*, *53*(1), 22-27.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (3rd ed.). Upper Saddle River, NJ: Pearson Education.

Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership, 68*(4), 35-39.

Danielson, C. (2012). Observing classroom practice. *Educational Leadership, 70*(5), 32-37.

Danielson, C. (2013). *Framework for teaching: Evaluation instrument*. Princeton, NJ: The Danielson Group.

Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional learning.* Princeton, NJ: Educational Testing Services.

Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education Policy Analysis Archives*, *1*(8), 1-44.

Darling-Hammond, L. (2010). *Evaluating teacher effectiveness: How teacher performance assessments can measure and improve teaching*. Retrieved from http://www.americanprogress.org/issues/education/report/2010/10/19/8502/evaluating-teacher-effectiveness

Donahue, E. (2013). *Teacher perceptions of evaluation impact on individual teaching practices*. Unpublished manuscript, Department of Statistical Research Methods, University of Northern Colorado, Greeley, CO.

Donaldson, M. L. (2009). *So long Lake Wobegon--Using teacher evaluation to raise teacher quality.* Retrieved from http://www.americanprogress.org/wp-content/uploads/issues/2009/06/pdf/teacher-evaluation

Duncan, A. (2009). *Partners in reform.* Retrieved from http://www2.ed.gov/print/news/speeches/2009/07/07022099.html

Educator Effectiveness Act, Senate Bill 10-19. Ensuring Quality Teaching Through Educator Effectiveness: Great Teachers and Leaders (2010) .

Eisner, E. W. (1991). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice.* New York: Macmillan.

Ellet, C. D., & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspective from the USA. *Journal of Personnel Evaluation in Education,17*(1), 101-128.

Glass, J. (2013). *Unparalleled altitude: A globally inspired vision for Eagle County schools*. Retrieved from http://issuu.com/eagleschools/docs/altitudereport3

Gordon, R., Kane, T. J., & Staiger, D. O. (2006). *The Hamilton project: Identifying effective teachers using performance on the job*. Washington, DC: The Brookings Institution.

Guba, E. G., & Lincoln, Y. S. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Fransisco, CA: Jossey-Bass.

Hamel, J., Dufour, S., & Fortin, D. (1993). *Case study methods.* Newbury Park, CA: Sage.

Heneman, H. G., & Milanowski, A. T. (2003). Continuing assessment of teacher reaction to standards-based teacher evaluation systems. *Journal of Personnel Evaluation in Education, 17*(2), 173-195.

Holtzapple, E. (2003). Criterion-related validity evidence for a standards-based teacher evaluation system. *Journal of Personnel Evaluation in Education*, *17*(3), 2007-219.

Jerald, C. (2012). *Movin' it and improvin' it--Using both education strategies to increase teaching effectiveness*. Retrieved from http://www.americanprogress.org/wp-content/uploads/issues/2012/01/pdf/movin_it_improvin_it.pdf

Johnson, S. M. (1990). *Teachers at work: Achieving success in our schools.* New York: Basic Books.

Johnson, S. M., & Fiarman, S. (2012). The potential of peer review. *Educational Leadership, 70*(3), 20-25.

Kagan, S., & Kagan, M. (2009). *Kagan cooperative learning.* San Clemente, CA: Kagan Publishing.

Kane, T., & Cantrell, S. (2012). *Learning about teaching: Initial findings from the measures for effective teaching project*. Seattle, WA: Bill and Melinda Gates Foundation.

Kane, T., Wooten, A. L., Taylor, E. S., & Tyler, J. H. (2011). *Evaluating teacher effectiveness*. Retrieved from http://educationnext.org/evaluating-teacher-effectiveness/

Kelley, C. (1999). The motivational impact of school-based performance awards. *Journal of Personnel Evaluation in Education, 12*(4), 309-326.

Kimball, S. M. (2002). Analysis of feedback, enabling conditions and fairness perceptions of teachers in three school districts with new standards-based evaluation systems. *Journal of Personnel Evaluation in Education, 16*(4), 241-268.

Kimball, S., White, B., Milanowski, A., & Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody Journal of Education, 79(*4), 54-78.

Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *American Journal of Occupational Therapy, 45*(3), 214-222.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* Newbury Park, CA: Sage.

Long, M. (2000). *Teacher effectiveness: Do teachers matter*. Retrieved from http://www.psych-ed.org?Topics/teacher_effectivenss.htm

McLaughlin, M. W., & Pfeifer, R. S. (1988). *Teacher evaluation: Improvement, accountability, and effective learning*. New York: Teachers College Press.

Marshall, K. (2005). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan, 86*(10), 727-735.

Marshall, K. (2012). Fine-tuning teacher evaluation. *Educational Leadership, 70*(3), 50-53.

Marzano, R. (2012). Two purposes of teacher evaluation. *Educational Leadership, 70*(3), 14-19.

Marzano, R. (2014). *Marzano teacher evaluation model.* Retrieved from http://www.marzanoevaluation.com/evaluation/four_domains/

Medley, D. M., & Coker, H. (1987). The accuracy of principal judgments of teacher performance. *Journal of Educational Research, 80*(4), 242-247.

Merriam, S. M. (2009). *Qualitative research: AgGuide to design and implementation*. San Francisco, CA: Jossey-Bass.

Meyer, J. P. (2010). Colorado teacher bill ignites firestorm of support, opposition. *The Denver Post*. Retrieved from http://www.denverpost.com/ci_14953971

Mielke, P., & Frontier, T. (2012). Keeping improvement in mind. *Educational Leadership, 70*(3), 10-13.

Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. *Peabody Journal of Education, 79*(4), 33-53.

Milanowski, A., Kimball, S., & Odden, A. (2005). Teacher accountability measures and links to learning. In L. Stiefel, A. E. Schwartz, R. Rubenstein, & J. Zabel (Eds.), *Measuring school performance and efficiency: Implications for practice and research* (pp. 137-159). Larchmont, NY: Eye on Education, Inc.

Moore-Johnson, S. (n. d.). *Why teachers must have an effective evaluation system*. Retrieved from http://www.aftma.net/educator-resources/teacher-evaluation

Morse, J. M., Barrett, M. Mayan, M. Olson, K., & Spiers, J. (2002). Verification strategies for establishing reliability and validity in qualitative research. *International Journal of Qualitative Methods. 1*(2), 1-19.

Namaghi, S. A. O. (2010). A data-driven conceptualization of teacher evaluation. *Qualitative Report 15*(6), 1504-1522.

National Commission on Teaching and America's Future. (1996). *What matters most: Teaching for Americas future.* Retrieved from http://eric.ed.gov/?id=ED395931

National Institute for Excellence in Teaching. (2014). TAP: *The system for teacher and student advancement*. Retrieved from http://http://www.talentedteachers.org

Odden, A. (2004). Lessons learned about standards-based teacher evaluation systems. *Peabody Journal of Education, 79*(4), 126-137.

Papay, J. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review, 82*(1), 13-141.

Patton, M. Q. (2002). *Qualitative research evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Peterson, K. D. (2000). *Teacher evaluation: A comprehensive guide to new direction and practices*. Thousand Oaks, CA: Corwin Press.

Phillips, V., & Weingarten, R. (2013). *Six steps to effective teacher development and evaluation.* Retrieved from http://www.newrepublic.com/article/112746/gates-foundation-sponsored-effective-teaching

Rubin, H. J., & Rubin, I. S. (1995). *Qualitative interviewing: The art of hearing data*. Thousand Oaks, CA: Sage

Schools Matter. (2010). *Colorado's SB 191*. Retrieved from http://www.schoolsmatter info/search?q=Senate+BIll+191

Shenton, A. K. (2004). *Strategies for ensuring trustworthiness in qualitative research projects*. Retrieved from http://www.crec.co.uk/docs/Trustworthypaper.pdf

Shields, C. M. (2007). *Can case studies achieve the "Gold Standard"? Or when methodology meets politics*. Paper presentation at the Annual Meeting of the American Educational Researcher Association, Chicago, IL.

Smyser, S., & Wodlinger, T. (2012). *The quality of the teacher has the greatest impact on student achievement*. Paper presented at Iowa Teacher Evaluation Forum, Des Moines, IA.

Sporte, S. E., Stevens, W. D., Healey, K., Jiang, J., & Hart, H. (2013). *Teacher evaluation in practice: Implementing Chicago's REACH students.* Retrieved from http://ccsr.uchicago.edu/publications/teacehr-evalautin-practice-implementing-chicagos-reach-students

Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lncoln (Eds.), *The Sage handbook of qualitative research* (3rd ed., pp. 443-466). Thousand Oaks, CA: Sage.

Stake, R. E. (2006). *Multiple case study analysis*. New York, NY: Guilford Press.

Stand Organization. (2010). *The passage of Colorado's Senate Bill 191: Game changing teacher and principal effectiveness legislation*. Retrieved from http://stand.org/sites/default/files/National/Case%20Study_THE%20PASSAGE%20OF%20COLORADO.pdf

Steele, J. L., Hamilton, L. S., & Stecher, B. M. (2010). *Incorporating student performance measures into teacher evaluation systems.* Retrieved from http://www.rand.org/pubs/technical_reports/TR917.html

Stiggins, R. J., & Duke, D. L. (1988). *The case for commitment to teacher growth: Research for teacher evaluation*. Albany, NY: State University of New York Press.

Stumbo, C., & McWalters, P. (2011). Measuring effectiveness: What will it take? *The Effective Educator, 68*(4), 10-15.

Taylor, E. S., & Tyler, J. H. (2012). Can teacher evaluation improve teaching--Evidence of systematic growth in mid-career teachers. *Education Next, 12*(4). Retrieved from http://www.educationnext.org/can-teacher-evaluation-improve-teaching/

The New Teacher Project. (2013). *Fixing classroom observations: Issue analysis report*. Retrieved from http://www.tntp.org/assets/documents/TNTP_FixingClassroomObservations_2013.pdf

Toch, T. (2008). Fixing teacher evaluation. *Educational Leadership, 66*(2), 32-37.

Toch, T., & Rothman, R. (2008). *Rush to judgment: Teacher evaluation in public education.* Washington, DC: Educational Sector.

U.S. Department of Education. (2009). *Race to the top program (executive summary).* Retrieved from http://www2.ed.gov/programs/racetothetop/index.html

U.S. Department of Education. (2010). *Great teachers and great leaders*. Retrieved from http://www2.ed.gov/policy/elsec/leg/blueprint/great-teachers-great-leaders.pdf

Varlas, L. (2012). Rethinking teacher evaluation. *ASCD Education Update 54*(12), 1-6.

Weglinsky, H. (2000). *How teaching matters: Bringing the classroom back into discussions of teacher quality*. Retrieved from https://www.ets.org/research/policy_research_reports/publications/report/2000/idxn

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect:-Our national failure to acknowledge and act on differences in teacher effectiveness.* Retrieved from http://files.eric.ed.gov/fulltext/ED515656.pdf

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*, 57-67.

Yin, R. K. (2009). *Case study research* (4th ed.). Thousand Oaks, CA: Sage.

APPENDIX A

DESCRIPTION OF THE EVALUATION SYSTEM UTILIZED IN MOUNTAIN SCHOOL DISTRICT

Mountain School District (MSD) adopted a system of teacher evaluation and professional growth more than ten years ago. Developed by the Milken Family Foundation, known as the Teacher Advancement Program, and currently managed by the National Institute for Excellence in Teaching (NIET; 2014), it is described as a system for teacher and student advancement. The system has evolved since its initial implementation but basic tenets including multiple career paths and ongoing professional development are still emphasized. System components include provisions for professional learning, observation, coaching, feedback, and modeling. Teacher input regarding changes in the language of the instructional evaluation rubric, methods of data collection, and strategies for providing coaching and feedback have been integrated into the existing system. The system also includes a component of performance pay. Although all of these components exist within the system currently utilized within Mountain School District, the focus of this study was on the structures for observation, coaching, feedback, and evaluation.

This system, now referred to as the MSD Teacher Evaluation System, provides multiple opportunities for differentiated teacher roles and responsibilities. Teacher roles are defined as career teacher, mentor teacher, or master teacher. Brief definitions are provided in Chapter II. A career teacher dedicates 100% of the instructional school day to direct contact with students while a mentor teacher and a master teacher divide time between students and teachers. The role of the mentor and master teacher includes coaching and modeling as well as an expectation to facilitate professional development during job-embedded Professional Learning Communities. The mentor teacher divides his or her time between classroom instruction and coaching in a 70% to 30% ratio. The

master teacher's time is the inverse, 30% classroom instruction and 70% teacher coaching as well as facilitation of professional development.

Mentor and master teachers are interviewed, rehired, or replaced on a one to three year cycle. While this cycle is recommended, it is not strictly enforced. Criteria that a mentor must meet in order to be considered for this position include:

- BA degree required.
- Relevant MA degree strongly preferred.
- Minimum of three years of successful teaching experience.
- Active teacher authorization or license issued by the State of Colorado.
- Must be "highly qualified" according to Colorado's definition of a "Highly Qualified" teacher under the No Child Left Behind Act (effective January 1, 2004 for current teachers and 2002-03 for newly hired teachers).
- Proven expertise in teacher professional development.
- Must be a certified evaluator in the Mountain School District (or obtain within 90 days of hire).
- Must be proficient in subject-matter content, instructional techniques and understand student needs.
- Excellent instruction, communication and interpersonal skills.
- Excellent time management and organizational skills.

For the master teacher position, the criteria are similar and include:

- MA degree and three years of successful teaching experience or a BA degree and seven years of successful teaching experience.
- Active teaching license issued by the State of Colorado.

- Must be "highly qualified" according to the Colorado's definition of a "Highly Qualified" teacher under the No Child Left Behind Act (effective January 1, 2004 for current teachers and 2002-03 for newly hired teachers).
- A minimum of 3 years of coaching or leadership experience.
- Must be recommended by his/her current principal.
- National Teaching Certification strongly preferred.
- Must be a certified evaluator in the Mountain School District (or obtain within 90 days of hire).
- Demonstrated expertise in content, curriculum development, student achievement, and assessment.
- Training on District wide initiatives (or obtain within six months of hire).
- Excellent communication and interpersonal skills.
- Excellent time management and organizational skills.

Within this system, value is placed on the evaluation rubric derived from the Danielson Frameworks for Teaching (Danielson & McGreal, 2000) as a coaching tool. Master and mentor teachers are trained annually to use the evaluation tool in order to identify areas of strength and weakness for teachers and to provide appropriate coaching and professional development opportunities. Special attention is focused on developing inter-rater reliability surrounding the implementation of the rubric. Throughout the school year, the evaluation rubric is used in an ongoing manner in order to develop a path for coaching and evaluation. Teachers experience three formal evaluations, including a post conference, per year. The mentor teacher, the master teacher, or the principal completes each formal observation and evaluation of a full-length lesson. In addition to

the formal observation and evaluation, teachers experience three walkthroughs, known as overtime observations, per evaluator. Walkthroughs are an attempt to respond to teacher input regarding the way that some discreet classroom instructional practices, though high quality, may not be realistically observed in every lesson. Teachers have three formal observations and nine walkthrough observations for a total of 12 classroom visits by trained evaluators yearly. Each of these 12 observations provides an opportunity for teachers to receive feedback on their practices.

APPENDIX B

INSTITUTIONAL REVIEW BOARD APPROVAL

Reply Reply All Forward

Fw: IRBNet Board Action

Donahue, Erika [dona2279@bears.unco.edu]

To: ERIKA DONAHUE

Tuesday, October 14, 2014 9:27 PM

From: Megan Babkes Stellino <no-reply@irbnet.org>
Sent: Monday, January 27, 2014 1:36 PM
To: Donahue, Erika
Subject: IRBNet Board Action

Please note that University of Northern Colorado (UNCO) IRB has taken the following action on IRBNet:

Project Title: [564765-1] Teacher Perceptions of the Impact of an Evaluation System on Classroom Instructional Practices
Principal Investigator: Erika Donahue

Submission Type: New Project
Date Submitted: January 25, 2014

Action: APPROVED
Effective Date: January 27, 2014
Review Type: Exempt Review

Should you have any questions you may contact Megan Babkes Stellino at megan.stellino@unco.edu.

Thank you,
The IRBNet Support Team

www.irbnet.org

APPENDIX C

CONSENT FORM FOR HUMAN PARTICIPANTS IN RESEARCH

UNIVERSITY of
NORTHERN COLORADO

CONSENT FORM FOR HUMAN PARTICIPANTS IN RESEARCH
UNIVERSITY OF NORTHERN COLORADO

Project Title: Teacher Perceptions of the Impact of an Evaluation System on Classroom Practice
Researcher: Erika Donahue, University of Northern Colorado doctoral student
Phone: 970-390-8687 E-mail: dona2279@bears.unco.edu
Research Advisors: Anthony Armenta, Ed.D., Linda Vogel, Ph.D.
Email: anthony.armenta@unco.edu, lrvogel_1122@yahoo.com

The purpose of this study is to determine teacher perceptions of how an existing evaluation system may inform change in instructional practices. While research exists which helps to identify the challenges inherent in teacher evaluation, very little is known about how the experience of being evaluated might change teacher effectiveness. Research has not specifically addressed how teachers perceive the systems they are bound within to influence their choice of classroom practices. The researcher hopes that this study will elucidate themes and topics that may lead to increased effectiveness of an existing evaluation system.

During a one-hour interview, the researcher will ask you your perceptions of the existing teacher evaluation system. Your involvement in this project will take place outside of school hours and school buildings at a date and time convenient for you. Questions addressed during the interview may include:

- Are there aspects of the teacher evaluation process that you perceive to have an impact on your teaching practices?
- Do you feel the existing system could be refined to better meet your needs for professional development? In what ways?

When the research has been completed, transcripts of your responses will be available to you, upon request. The researcher will take precautions in order to protect the confidentiality of your responses, as you will be asked to select a pseudonym or have a pseudonym assigned to you. No district employees, other than the researcher, will know of your involvement unless you choose to share your experience. All interview data and transcripts will be stored on an external hard drive, in a locked cabinet, and the researcher will be the only individual with access to these data. This information will be destroyed three years after the completion of this research study. Copies of consent forms will be stored in the locked office of the research advisor. These will also be destroyed three years after the completion of the study.

The potential risks and discomforts to you are minimal, as the responses you provide to the interview questions will not be attributed to you as an individual. You are invited to participate and may agree to this participation through email consent. You have the ability to choose not to participate and will have the option to end your participation in the research at any time.

The potential benefits of this research include the possible use of your contributions in the improvement of the existing teacher evaluation system.

Costs for you are minimal. Interviews will take place at a location and time of your choosing. The researcher will provide drinks, snacks, and a $10 gift card from a local bookseller to thank you for your time. You are responsible for meeting your own travel expenses if you select a venue for the interview requiring travel.

Participation is voluntary. You may decide not to participate in this study and if you begin participation you may still decide to stop and withdraw at any time. Your decision will be respected and will not result in loss of benefits to which you are otherwise entitled. Having read the above and having had an opportunity to ask any questions, please sign below if you would like to participate in this research. A copy of this form will be given to you to retain for future reference. If you have any concerns about your selection or treatment as a research participant, please contact the Office of Sponsored Programs, Kepner Hall, University of Northern Colorado Greeley, CO 80639; 970-351-2161.

Participant's Signature	Date
Researcher's Signature	Date

APPENDIX D

CONTEXT BUILDING INTERVIEW QUESTIONS

Context Building Interview Questions

1. What grade level, content area, or subject do you currently teach? How long have you taught at this level?
2. Please tell me about your background as a teacher.
3. Have you ever been a mentor or master teacher within MSD?
4. Have you ever participated in professional development or received training in the evaluation system or evaluation rubric?
5. In terms of evaluation scores, would you place your final rating from the 2012-2013 school year within the scoring range of 1.5-1.9, 2.0-2.9, 3.0-3.9, or 4.0-4.9?

APPENDIX E

INTERVIEW QUESTIONS RELATED TO THE PROPOSITIONS OF THE STUDY

Interview Questions Related to the Propositions of the Study

1. What do you think improves teaching practices?
2. How do you receive feedback on your teaching?
3. Have you had an experience where you received feedback and tried to implement changes in practice?
4. Are there aspects of the existing teacher evaluation process that you perceive to have an impact on your teaching practice?
5. Are there aspects of the evaluation system that you feel prohibit you from changing your practice?
6. What would you like to share about a specific experience you have had with the evaluation process in this school district?
7. How do you feel about the way the evaluation tool measures your teaching?
8. In your experience, how have you seen changes in your evaluation results connected to changes in your teaching practices?
9. What are the essential characteristics that teacher evaluation must possess in order for teachers to make the most of an evaluation experience?
10. Do you feel the existing evaluation system could be refined to better meet your need for professional growth? In what ways?
11. What else would you like me to know or understand about your experience with teacher evaluation?
12. How do you feel about the performance pay element of the evaluation system in MSD?
13. How does the following statement resonate with you: 'Teachers want to be great but don't know how or what to do to get better.'